UNIVERSITETET I AGDER

The relationship between default on loans and operating costs in microfinance institutions

MARIT KRINGLEN

SUPERVISOR
Roy Mersland

University of Agder, 2016
School of Business and Law

Preface and Acknowledgements

This study is conducted as a finishing step in completing my Masters in Business Administration at the University of Agder. The study accounts for 30 credit points and is carried out through one semester.

A huge motivation in my work on this thesis is that to my knowledge, no similar study has previously been conducted. It has made the process very interesting and exciting, but also very challenging. The process has allowed me to test the skills acquired through the five-year masters program, while simultaneously being very educational. Writing the master thesis has taught me a lot about conducting research and has enlightened the possibilities and challenges that comes with it.

The research topic was suggested to my by supervisor Roy Mersland, who has great experience researching microfinance institutions. The relationship between the default rate and operating costs rate stroke me as very interesting and I was surprised to find that no one had previously studied this in the microfinance industry. It has truly been an exciting and educating process, and I am thankful for the opportunity to work on this subject.

I would like to thank Roy Mersland for his contribution on the research topic and his guidance throughout the process. I would further like to thank Stephen Zamore, doctorate candidate at the University of Agder, for great advice and guidance in my study, and especially his support and great knowledge in statistics. Finally I want to thank family and friends who have supported me throughout my education and this final process. I would not have been able to do this without any of them.

Marit Kringlen
Kristiansand, 3. October 2016

Abstract

One of the challenges in microfinance is to solve asymmetric information and cost concerns related to serving poor customers with little or no collateral, in order to offer poor people and small businesses access to financial services. In resent times the industry has been critiqued for maintaining harsh collection practices and charging too high interest rates to their clients. The microfinance mission of serving the poor has been questioned from several holds. There are sufficient previous studies that show that the interest rate to a great extent is driven by the operating costs in the institutions. Identifying the drivers of the operating costs thus becomes necessary in order to lower interest rates offered to the customers. The critique on collection practices is not uncalled for. Frequent collection of repayment is in the microfinance industry generally viewed as an essential component in reducing the risk of default, which historically has been a prominent goal of microfinance institutions due to the lack of collateral offered by their clients. The lack of credit history further adds to the need to manage risk, but with little background information allocation of resources becomes problematic. Transaction costs are therefore high in the industry.

It is a paradox that the world's poorest are charged with the highest cost of capital, and the industry faces a need for lower interest rates in order to help more people. This study investigates the relationship between default rates and operating costs in the microfinance sector, and looks into whether the microfinance sector over time has been too concerned with lowering default rates. Has the focus on default rates left the microfinance institutions with too high operating costs due to extensive transaction costs connected to monitoring, control and collection practices? Would they be better off by allowing for higher levels of default and lower transaction costs?

List of tables

List of figures

Table of Contents

1. Introduction

This chapter discusses the background for the study of the relationship between the default loans rates and the operating costs rate in microfinance institutions. It provides a statement of the research problem, the research objective and the research question, as well as the contribution and organisation of the study.

1.1 Background of the study

Microfinance institutions are organisations that offer banking services to poor people who are economically active and with a need to loan small amounts to finance for example a business idea or education, or to manage emergencies, obtain assets or smooth consumption (Christen, Lyman & Rosenberg, 2002). These clients often lack credit histories and/or collateral, and therefore experience difficulties accessing financing from ordinary commercial banks (Banerjee & Duflo, 2007). Microfinance institutions are therefore often viewed as contributors in creating economic opportunities and poverty alleviation (Di Bella, 2011).

Existing literature on microfinance institutions' operating costs rate focus on, amongst others, the type of ownership (see Mersland & Strøm, 2008; Mersland, 2009), characteristics of the microfinance institutions (see Gonzales, 2007) and whether or not the institutions are receiving subsidies (see Caudill, Gropper & Hartarska, 2009). To my knowledge, there are no published empirical studies related to the relationship between the level of default on loans and the operating costs rate of the institutions. This study offers to close the gap by examining the nature of this relationship and the direction of its effects.

1.2 Statement of the research problem

Microfinance institutions serve a social mission of outreach. By offering small loans they are able to extend credit to more people, including those who are only able to repay very small amounts, i.e. the poor people. Because microfinance customers often offer little or no collateral, the risk inferred is generally higher than for mainstream commercial banks. Over time, this has brought with it a need to lower risk by spending great resources on reducing the level of default on loans. This thesis is based on the idea that as high levels of default evidently will lead to increased costs, very low levels may similarly increase the operating costs, as the manpower demanded to maintain these levels is costly. As will be argued in later chapters, high operating costs are assumed to affect the interest rate levels offered to customers. The mission to help the poor in establishing businesses or smooth consumption can more easily be fulfilled if the interest rates offered to them are more affordable. Therefore, an efficiency analysis focused on the effect of the default rate on operating costs is timely and important.

1.3 Research objective

The research objective of this thesis is to study the relationship between the default on loans rate and operating costs rate in microfinance institutions.

1.4 Research question

What is the relationship between the default on loans rate and operating costs rate in microfinance institutions?

1.5 Contribution of the study

Building on contract theory, principle-agent and moral hazard theory, theory on credit risk and Berger & De Young's (1997) research on the relationship between problem loans and cost efficiency in commercial banks, this study aims to provide a better understanding of this relationship in microfinance institutions. It uses multivariate analysis with instrumental variables to control for the direction of the relationship and to produce a graphic illustration of it. The study adds to existing literature within microfinance by suggesting that adjusting the default on loans rate in either direction could potentially alter the cost efficiency in the industry.

1.6 Organisation of the study

The thesis is divided into seven chapters. Following the introduction, which reviews the background of the study and the research objective, the second chapter focuses on the microfinance industry and deliberate on the motivation for the study. Theories and past research is presented in chapter three, along with the outline of the conceptual framework and hypothesis. Chapter four is concerned with the data and offers information about the sample and an outline of the variables. Following the fourth chapter, the research methodology is discussed in chapter five, and the empirical findings and results in chapter six. Chapter seven rounds up the study by presenting the main conclusions and suggestions for further research.

2. The microfinance industry

This chapter provides an overview of the history and development of microfinance, the mission of microfinance, and the drivers of the microfinance lending rate.

2.1 Microfinance history and development

Microfinance is by Helms (2006) defined as the supply of banking services to poor families and micro enterprises. Microfinance institutions usually offer small loans of short duration without formal collateral, often set up as group loans for which the group members are jointly liable for repayment. The concept was developed in the 1970s and 1980s (Mersland & Strøm, 2012a) as a reaction to the discouraged development resulting from subsidized rural credit in the two prior decades (Adams & Finchett, 1992). The 1950s and 1960s were characterized by international donors and national governments investing vast amounts in low-cost credit to farmers, ultimately resulting in intensification of corruption and high default rates (Hulme and Mosely 1996).

The innovation that came with microfinance in the 1970s and 1980s focused on aligning poor people's demand for financial services with the requirement of repayment to the banks. The contemporary way of organising the lending agreement had its roots in informal financial systems, like that of rotating savings and credit associations (Adams and Fitchett, 1992), where poor people come together to organize small credit schemes and savings clubs. Banks ensured repayment by issuing only small loans of short duration, and would back these up with informal or group collateral (Mersland & Strøm, 2012a). Further, in contrast to the practise in the 1950s and 1960s, the microfinance institutions would assess the payment capacity of the customers based on their current sources of income rather than on anticipated income from new business.

The start of the microfinance era is by many people associated with Bangladeshi Mohammad Yunus, who in 1976 began issuing small private loans to women after a visit to the poorest parts of the village Jobra, in India (Yunus & Jolis, 1999). He noticed that very small loans would allow for disproportionate changes and opportunities to these people, and was motivated to build on the possibility to make a difference (Yunus & Jolis, 1999). Yunus later developed a relationship with the governmentally regulated Janata Bank to secure loans to the people of Jobra. Soon after, in the early 1980s, the Grameen Bank was established with the sole purpose of providing financial service to the poor people in the village (Yunus & Jolis, 1999). This is perhaps the best-known microfinance institution from the early years. Despite being the best-known, it was not the first. Opportunity International, one of today's the biggest international microfinance networks, roots back to 1971 when pioneer David Bussau and Al Wittaker started issuing small loans to engender work for the poor people in the area (Opportunity International, 2016).

Until the early 1990s microfinance initiatives were for the most part driven by donor-funded non-government organisations providing credit to entrepreneurial poor people (Mersland & Strøm, 2012a). However, as the years went on, operations developed to include all types of financial services. Today saving, insurance and systems for money transfer are common services offered by the microfinance institutions, and they are becoming available for the poor all over the world (Christen, Rosenberg & Jayadeva, 2004). The microfinance industry has in the recent decades seen a remarkable development, with microfinance institutions as the providers of the microfinance services. (Mersland & Strøm, 2013) report that the growth in the total loan portfolio of the microfinance institutions has been positive for the past two decades, and that the growth has averaged between 40% and 60% for several years. Also the growth rate in individual microfinance institutions has been strong, averaging over 20% annually for several years.

However, in the past ten years the maturation of the microfinance industry has brought with it claims that the industry is deserting the mission to serve the poor (Dichter & Harper, 2007). The media coverage concerning microfinance shifted rather rapidly from being praising and rosy in 2005 and 2006 to rather critical and grim in 2007. In 2005 the United Nations declared the year as the Year of Microcredit (United Nations, 2016), with Secretary-General Kofi Annan stating that: "microfinance has proved its value, in many countries, as a weapon against poverty and hunger. It really can change peoples' lives for the better, especially the lives of those who need it most" (United Nations, 2016). The glory continued into 2006 when microfinance pioneer Mohammad Yunus and the Grameen Bank was awarded the Nobel Peace Prize (Nobel Media, 2016). In 2007, however, the focus shifted to cover the Banco Compartamos case in Mexico, where the sale of overpriced loans received much criticism and changed the way the public viewed the industry (Rosenberg, 2007). Later, the tragic case of suicides in Andhra Pradesh continued to attract public attention and provoke scepticism about the microfinance concept (Ryhne, 2011). The case refers to several borrowers who committed suicide because they were unable to repay debt to microfinance institutions and local moneylenders (Business Insider, 2012, 24.2.). In the aftermath, accusations that microfinance institutions are overcharging interest, have drifted from their mission statement to help the poor, and are too hard-handed in collecting repayments on loans started to flourish. As a result, attention is now given to limit the size of the lending rate (Mersland & Strøm, 2012b).

Regardless of the concerns and scepticism, the growth in the industry has persisted in the following years. Mersland & Strøm (2012a) show that even in 2008, the year of the financial crisis, the growth was positive, though less than in previous and preceding years. Furthermore, their research indicate that a consolidation is under way in the industry, and that the average loan portfolio among the microfinance institutions were in 2009 over twice the size of that in 2007. Such consolidation illustrates that the amount lent to poor customers improve even when the number of institutions decreases. Naturally, this tremendous growth cannot take place unless it is advantageous for borrowers to undertake the loans, and Mersland & Strøm (2013) argue that the high levels of growth are evidence of the lending rate being acceptable.

2.2 Microfinance mission

A common trait in microfinance is the ability to solve asymmetric information and cost concerns related to serving poor customers with little or no collateral (Karlan & Zinman, 2009), thus giving unfortunate people and small businesses access to financial services. It targets poor people and small businesses in developing countries, and the microfinance institutions often further specify their target markets to people in semi-urban and rural districts and women (Mersland & Strøm, 2012a). Since its emergence, microfinance has kept entrepreneurial poor people as its main target, and has received criticism (see e.g. Helms, 2006) for thus failing to reach people with little or no entrepreneurial activities, who often are also the poorest people. However, the strong and persistent growth in the industry must arguably reflect a response to an underlying demand, and is thus an indication of microfinance having positive impact.

Providers of microfinance typically have both financial and social objectives (Armendariz & Morduch, 2010), that is, they have a double objective to provide financial services to the poor and to do so in a financially sustainable way. The idea of microfinance is to bring basic utility of finance to poor people (Green, Kirkpatrick & Murinde 2005), by providing people with the chance to smooth consumption and store savings. In this prospect, microfinance could be seen chance to extend financial services to people who previously have not had such opportunities. By removing the frictions that prevent poorer segments from access to financial services, Mersland and Strøm (2012a) suggest that the development of the county's financial system could improve. Morduch (1999) claims that access to microfinance while paying for the services can be seen as a tool to reduce poverty, while Levine (2005) shows that financial development has an influential effect on economic growth as well as income inequality. Recently, however, it has been questioned whether the endowment of small loans is the best solution to help poor people out of poverty. Further criticism has been raised on the high level of interest rates in the industry, with arguments that microfinance institutions only operates to earn money and are too attentive when it comes to obtaining repayments on loans (Mersland & Strøm, 2012b). The social objective of microfinance institutions fosters a need for lower interest rates by the simple justification that lower rates will make the loans more affordable and thus more available to the customers, serving the social objective of helping more people.

2.3 The drivers of the microfinance lending rate

In the light of the negative attention attracted to the industry in connection with the Banco Compartamos and Andhra Pradesh case accusations that microfinance institutions are overcharging interest started to flourish. As a result, attention is now given to limit the size of the lending rate (Mersland & Strøm, 2012b). Bhatt (2001) points out that microfinance institutions in many cases have experienced problems with high default rates. Defaults are bad for the industry not just because of the losses the microfinance institutions incur, but also of political and social concerns (Chakrabarty & Bass, 2013). Inability to meet financial obligations has proven to be a trigger for serious social effects such as riots, deterioration of community relationships and even suicide and death (Hulme, 2000; Montgomery, 1996). Chakrabarty & Bass (2013) find that this is especially observable in emerging markets, where political, social and economic risks are high in general, leaving doing business more difficult. Because these markets are more risky to operate in, microfinance institutions make an extra effort to try limit the risk they undertake. Field & Pande (2008) states that frequent collection of repayment installments is commonly believed to be one of the most important components in reducing the risk of default. Thus, great resources are often spent in the collection and monitoring process, and the transaction costs of the microfinance institutions are thereby driven up (Field & Pande, 2008). In order for microfinance institutions to be financially sustainable the high operating costs are further passed on to borrowers in terms of high interest rates ((Dehejia, Montgomery, & Morduch, 2012; Fernando, 2006; Morduch, 2000). Additionally, in emerging markets, which often are those that microfinance institutions operate in, inefficient litigation in dysfunctional courts are common, making contracts difficult to enforce. As such, it can be argued that borrowers will not worry about breaching contracts (Field & Pande, 2008). Thus, loan defaults pose potentially great risks for microfinance institutions.

Rosenberg, Gaul, Ford & Tomilova (2013) maintain that the interest rates are often much higher in the microfinance industry compared to regular banks, primarily because it is much more expensive to lend and collect on many small loans relative to fewer and larger loans. They too enlighten that the costs have to be covered through the interest rates charged to the customers. Furthermore, Rosenberg, Gonzalez & Narain (2009) reveals that the operating costs in microfinance institutions justify more than 50% of these rates. Gonzalez (2007) declare microfinance a high touch, high cost industry, and assert that identifying the drivers of the operating costs are necessary in order to lower interest rates offered to the customers. A lowering of the interest rate could in turn help microfinance institutions better reach their social mission of outreach to the poor. It is a paradox that the world's poorest are charged with the highest cost of capital. Offering small loans at affordable costs becomes one of the core challenges of the industry.

The motivation for investigating the operating costs in microfinance institutions thus rest on findings that the high interest rates in the industry are mainly caused by the operating costs of the banks (Gonzalez, 2010; Mersland & Strøm, 2012b; Rosenberg et al.,2013). Mersland & Strøm (2012b) find that for most microfinance institutions modifications in lending rates, revenue and profitability are products of increased input prices. They further hold that high costs and low margins is the industry's main problem. That statement is supported in Mersland & Strøm (2013), which finds that contrary to being an industry with high profits, it struggles with high costs and low earnings. This is maintained in our data set, where we see that the average portfolio yield is close to 38%, whereas the operating costs of portfolio is above 30% (see chapter 4.1). In addition to operating costs, the cost of funds and the loan loss have to be covered before profits can be distributed. Mersland & Strøm (2012b) report similar numbers and concludes that the average profitability in the microfinance sector is low and that return on assets often comes close to zero. Chapter 4.1 of this thesis will display similar findings in our dataset.

3. Theories and past research

This chapter discusses the empirical evidence from past research on default on loans, operating costs and social outreach in microfinance institutions, and presents existing literature on the relationship between cost efficiency and problem loans.
It will look into the nature of this relationship in search for an enhanced understanding of the influence of the loan default rate on the microfinance operating costs, and ultimately, the lending rate.

Further chapters will continue into investigating whether there might be a general functional form for this relationship that can further be used in search for an optimal level of default. Literature is drawn from microfinance and economical/financial studies and theory. The basis of the research question builds upon contract theory, agency theory and theories on moral hazard and credit risk, as well as the findings of Berger & De Young (1997) of reversed causality between problem loans and costs in banks. Although the Berger and De Young base their research on commercial banks, this paper will argue that the findings are likely to be similar when looked at from microfinance perspective. Most microfinance research is found in development journals, but in the later years we see more and more articles published in economic and financial journals as well. Yet the research available in mainstream financial journals is still limited. The chapter will be rounded of by an overview of the conceptual framework and a presentation of the hypothesis.

3.1 Theoretical background

3.1.1 Contract theory

Contract theory examines how contractual measures and agreements can be composed in the presence of asymmetric information, and is concerned with theories of incentives, information and economic institutions (Bolton & Dewatripont, 2005). The concept is widely discussed in microeconomics (see e.g. Hart & Holmström, 1987; Tirole, 1988; Maskin & Tirole, 1992; Dewatripont & Maskin, 1995; Crémer, Khalil &

Rochet, 1998; Bolton & Dewatripont, 2005; Köszegi, 2014), and while the literature covers several viewpoints and aspects of contract theory, it is common to use numerical utility structures to present the behaviour of decision-makers before implementing an algorithm for optimization. Contract theory discuss, amongst other, theoretical ways to deal with principle-agent issues such as moral hazard, adverse selection and signalling through the use of mathematical characteristics of the utility structure between the principal and the agent. In summary, contract theory is concerned with the economic analysis of contracts.

3.1.2 Agency theory

Agency theory has its roots in risk sharing literature, which is concerned with cooperating parties who have different attitudes towards risk (Eisenhardt, 1989), but expands to include different goals and labour division between the parties (Jensen & Meckling, 1976). More specifically, agency theory is concerned with solving the principal-agent challenges that can occur when tasks or responsibilities are delegated by one party (the principal) to another (the agent). Consequently, principal-agent problems typically occur in situations where both parties are utility maximizing, i.e. when the agent is concerned with maximizing his own utility, even when it might be on the expense of that of the principal (Fama & Jensen, 1983). Meckling & Jensen (1976) describe how agency theory depicts this relationship by using the metaphor of a contract. Building on the economic concepts of game theory and rational choice theory, agency theory focuses on how to create efficient contracts to govern the principal-agent relationship, given the available information (often limited and/or asymmetric) and goals, attitudes and incentives of the parties (MacNeil, 2000; Bromiley, 2005). Principal-agent theory is often exemplified using shareholder-manager relationships, but can be attributed to any situation where one party acts on behalf of another.

Bruce, Buck & Main (2005) argue that agency theory is based on the assumption that agents are self-interested and utility maximizing, which in many relationships may not be the case. Critics further contend that constraints external to the principle-agent relationship may limit the opportunistic behaviours of the agent or interrupt the systems used in monitoring and controlling agent behaviour (see e.g. Fligstein &

Freeland, 1995; Aguilera & Jackson, 2003; Bruce et al., 2005). Wiseman, Cuevas-Rodríguez & Gomez-Mejia (2012) assert that such constraints are specific to the institutional environment of a firm. The microfinance industry in this context, stand out from commercial banking in their innovative methods of social monitoring via group lending arrangements. The idea is that the environment in which the microfinance clients find them self will not allow for opportunistic behaviour by the borrower (i.e. behaviour that is not aligned with the agreements made with the lender, such as not meeting their obligations or withholding information about the success of operations), as the group members are jointly reliable for repayment of the loans. Thus, microfinance institutions rely on the agents (i.e. the microfinance customers) to use social pressure to avoid defaults that negatively affect the whole group. Nevertheless, there exists extensive literature supporting the central idea of agency theory; agents left unmonitored are likely to pursue private interests that deviate and even conflict with the goals of the principal (see e.g. Tosi, Werner, Katz & Gomez-Mejia, 2000; Westphal & Khanna, 2003; Harris & Bromiley, 2007). One could argue that this sort of behaviour is even more likely in the microfinance industry, as the clients often are poor and desperate. Additionally, the loans are regularly given with little or no collateral and with little information about the borrower and his/her credit history. It is reasonable to envision that without a certain degree of monitoring (formal or informal), and because credit history is poorly documented in the industry, microfinance customers could fairly easily abandon their agreement with the microfinance institution and walk out on their obligations in terms of repayment. It becomes apparent that microfinance institutions are prone to great risk, which is likely the main reason to why the industry over the past decades has put in such extensive resources in collection practice and monitoring to reduce it.

3.1.3 Moral hazard

Moral hazard is a widely discussed phenomenon within the agency theory (Milgrom & Roberts, 1992). It describes situations in which the one making decisions about risk is not the one responsible for the outcome (Krugman, 2009). This often occurs in situations where information asymmetry is present, i.e. that the decision-maker knows more about its actions or intentions than the party bearing the consequences of the risk, and when incentives varies between the two parties.

The term “moral hazard” was established in the 1600’s and used about immoral behaviour, often related to insurance relations (Dembe & Boden, 2000). Dembe & Boden (2000) do however explain that the term was renewed in the 1900’s and is now used to describe inefficiencies as a result of information asymmetry, rather than the morale of the involved parties. As implied by Dembe og Boden (2000), the concept of moral hazard has its roots in the insurance industry, which is built upon the idea of transferring risk to another party (Pritchard, 2016). The theory of moral hazard in that context is that the insurance taker will act differently knowing that he is insured than he would if he carried to full risk of costs himself. For example, knowing that they are insured, he may be less careful with his assets. The same logic can be transferred into the relationship between a lender and a borrower. After being granted a loan, the borrower may act recklessly or invest or spend money in a different way than the lender would prefer or that is agreed upon in their contract. In other words, principle-agent concerns arise when incentives between the borrower and lender do not align.

Existing contract-theoretic literature focus on how moral hazard evade the first-best solution, which is the one that would be obtained under complete information (see Hart & Homström, 1987; Rogerson, 1985; Schmitz, 2005). The literature discusses two main reasons, where the first assumes that the agent is risk-averse and the second assumes that the agent is risk-neutral, but wealth-constrained. In the first case the principle faces a trade-off between presenting the agent with incentives and insurance, whereas in the second case the agent (borrower) might experience problems repaying the principle (lender) so that there is a trade-off between the providing the agent with incentives and minimizing his limited-liability rent (Laffont & Martimort, 2002). In the microfinance industry, the institutions have to face and deal with both scenarios. The second case, where the agent (i.e. the microfinance client) is wealth-constraint, marks the foundation of the entire industry, which is centred at providing financial services to the poor. The microfinance institutions therefore have to constantly maintain a balance between offering incentives and reassurances to the borrowers, while exercising monitoring and control measures in order to reduce risk of default. Armendariz de Aghion & Morduch (2005) explain that microfinance institutions are exposed to moral hazard from their credit clients because they do not have sufficient information about the borrowers to separate good from bad risk. The solution has therefore in many cases been to apply the same level of effort in collection of repayment and control and monitoring to all customers. This of course, is very costly in terms of use of recourses. The moral hazard concern is in microfinance more apparent than in commercial banks, as microfinance customers often have little or no existing credit history documented (Banerjee & Duflo, 2007), which in turn set the basis for information asymmetry and adverse selection.

3.1.4 Credit risk

Credit risk refers to a financial institution's likelihood of loss due to a borrower's default on debt (SAS, 2016). The loss refers to, amongst others, unpaid principal and interest, interference with cash flows and augmented collection costs. Consequently, greater levels of credit risk are associated with elevated costs of lending (operating costs) (Simkovic, 2016). Credit risk management is the measures taken to deal with this risk (Basel Committee on Banking Supervision, 2000), and involves identification, measurement, monitoring and control of the risk that arise from a possible default (Coyle, 2000). The Basel Committee on Banking Supervision (2000) state that effective managing of credit risk is fundamental to the long-term achievements of financial institutions, and that the loan portfolio constitutes the main component of credit risk. Measuring the credit risk is however not so straight forward, due to the many factors that may influence the borrowers ability to repay the lender (Investing Answers, 2016). Some suggested sources to credit risk in microfinance institutions are offered below:

- Microfinance customers may struggle to generate sufficient return on investments due to lack of knowledge and restricted access to technical advice or support services.
- Insufficient return on investment may be caused by political challenges or natural disasters.
- Lack of credit history and prevalent use of informal financing in regions where microfinance institutions operate pose the risk that the microfinance customer may face liabilities towards informal lenders. These may get precedence over the microfinance institution due to the often-ill consequences offered by the informal lenders.

- Unexpected circumstances at the borrowers household, such as sudden illness, accident or death (pure risk) may interrupt business activities or lead to redirected attention and use of funds.
- Microfinance customers may redirect the use of funds to non-essential consumption, often in connection with lacking incentives for repayment and/or lack of motivation in terms of generating cash flows through investments/entrepreneurial activities.

The fundamental concepts of credit risk management have been portrayed by numerous authors such as Lindergren (1987), Santomero and Babbel (1997), Dowd, Bartlett, Chaplin, and Kelliher & O'Brien (2008) as: "(i) the establishment of a clear risk policy and a reporting structure; (ii) underwriting authority and loans limit; (iii) allocation of responsibility and accountability; (iv) prioritization of the lending process and systems; and (v) the timely communication of risk information to top management" (Afriyie & Akotey, 2013). McKinsey&Company (2016) assert that a well-designed credit process can reduce a business' operating expenses by 15-20%, and that financial institutions must have a hands-on approach in handling possible losses to sustain value. SAS (2016) further maintain that the starting point to achieving effective credit risk management is to acquire a complete understanding of a bank's overall credit risk by screening risk at the individual, customer and portfolio levels. This can however be challenging in microfinance institutions, as information is often asymmetric and credit history poorly documented (Armendariz de Aghion & Morduch (2005); Banerjee & Duflo, 2007). While banks strive for a thorough outline of their risk profiles, sufficient and correct information can be hard to obtain in an environment characterized as "unbankable" by commercial financial institutions. Without a thorough risk assessment, institutions struggle to assess whether their capital reserves accurately reflect risks or if loan loss reserves sufficiently cover possible short-term credit losses (SAS, 2016). A possible way to deal with the restricted access to credit risk assessment is to reduce risk all together by tightening in on monitoring and supervision of customers. This is likely what the microfinance industry has experienced over the past few decades, which has resulted in very low levels of default in the industry (Field & Pande, 2008). Because the institutions are not able to make accurate assessments of the likelihood of default, resources are directed at collection practice and supervision of all the microfinance customers. Whereas commercial banks with well-established credit risk management are able to sort out where to allocate resources to reduce the risk of default, microfinance institutions struggle to make the distinction between good and bad risk. It becomes apparent that the microfinance industry compensate for the lacking advantages of effective credit risk management by appointing great resources to avoid default in “all” cases. To investigate whether the strong focus on default rates demonstrates the best use of resources is both timely and necessary.

On the other hand, handling (i.e. limiting) the risk of default is important for any financial institution, as defaults have been shown to leave banks with fewer resources available to for lending to other customers, deflate staff moral and affect borrower confidence (Agu, 1998). Research also shows that the operational costs associated with loans past due tends to be extensive and reduce the profitability of the banks (Padmanabham, 1988; Agu, 1998). Inadequately managed risk can further result in stakeholders (e.g. investors, customers, lenders, donors, savers and staff) losing confidence in the bank, and ultimately lead to reduced access to funding (Natilson & Bruett, 2001). Without, or with limited access to, funding microfinance institutions will struggle to meet their social objective of providing financial services to the poor. As pointed out by several authors (see Morduch, 1999; Sebstad & Cohen, 2000; Levine, 2005; Green, Kirkpatrick & Murinde, 2005), access to financial services through appropriate delivery mechanisms can help microfinance clients reduce their vulnerability and improve their life quality. This provides a fundamental basis for reducing poverty as well as strengthening the sustainability of the microfinance institution.

3.1.5 Cost efficiency

Berger & De Young (1997) is widely known and well credited for their studies on the relationship between problem loans and cost efficiency in banking institutions. Their main findings show that banks incur additional costs from loans that do not perform or default because they force the bank to spend extra resources on monitoring and underwriting to influence loan quality. They do however argue that whether or not the default on loans rate can be included as part of operating costs depends on the nature of the relationship between these two measurements. Their research then investigates this relationship by examining several of the banks' policies and underlying concerns, such as the cause of defaults and bank failures and the supervisory focus of bank managers. Their main argument is that when loans become non-accruing managerial effort is commonly upscaled to deal with these predicaments, thus leading to increasing expenses generally recognized as operating costs. These include, but are not limited to costs associated with monitoring borrowers and the value of their collateral, analysing and negotiating new repayment plans, acting on their right to collect (and dispose of) collateral, effort to maintain bank status and reputation, effort to avoid issues with loans that are currently performing, and shifts in management focus to problem loans on the expense of other tasks (Berger & De Young, 1997). Apart from collecting, maintaining and disposing of collateral, microfinance institutions face most of the same costs associated with problem loans and defaults, as their operations in that regard to a great extent is organized similarly to regular banks. One could even argue that the costs in the microfinance industry is even greater, do to lack of formal channels in which contracts can be enforced.

The findings in Berger & De Young (1997) support their expectation of a positive relationship between loans that are past due and operating costs, but the study shows that the effects are typically very small (approximately 1.7% decrease in cost efficiency as a consequence of a 1% increase in non-performing loans). They do however point out individual differences between banks and argue that the effects may be greater when the changes are larger (non-linear relationship). On the other hand, their study also show that bad management in terms of monitoring and underwriting (i.e. bad credit risk management) will lead to increased operating costs almost immediately, whereas loan defaults typically occur at a later point in time. This indicates that there may be challenges with endogeneity, as they do find evidence that after banks experience a decrease in cost efficiency, the level of non-accruing loans increase. Thus, the relationship between default on loans and operating costs seem to run in both directions. Although past research mainly has been based on the assumption of a positive relationship between default on loans and operating costs, the idea of default on loans negatively influencing the operating cost rate has been entertained in later research, e.g. in Mersland & Strøm (2013) where they note that a higher default rate may reduce operational costs if the case is that the microfinance institution is appointing great effort in obtaining repayment on loans. There is however a lack of studies to support this statement. This thesis looks to close this gap somewhat by investigating the relationship between default on loans and operating costs in microfinance institutions. The study will account for the suspected endogenity issues by including an instrument variable to the model, thus adjusting for reversed causality. Consequently, the thesis studies the order of the relationship between the default on loans rate and operating costs rate as well as the direction of it.

3.2 Past research

Bateman (2010) has become one of the more known critics of microfinance with his claims that it does not work because of lack of consideration for people's well-being. This claim is based on critique that microfinance institutions no longer follow their social mission of outreach to the poor. He further asserts that rather than helping clients smooth consumption and overcome poverty, microfinance institutions are chasing high profits and returns. His work provides an overview of problem areas in the industry that must be recognized, but there are however few studies to support his claims of mission drift, especially when viewing the industry as a whole. His critique does nevertheless raise important issues such as whether the level of the interest rate offered to microfinance customers is too high and if microfinance institutions operate with too high profitability. Mersland & Strøm (2013) investigate these questions from a cost perspective and find that the lending rate is strongly clustered around the zero profit margin and that the trend line is basically flat, starting with a zero profit margin and following a weak negative inclination. This shows that for their sample of 405 microfinance institutions in 73 countries the generally high lending rates are due to other causes than high profit margins. In fact, their study suggests that microfinance institutions in general are reinforcing their position in the poorest client segments as they age, indicating that they are not operating on a profit motive.

Conversely, Rhyne (1998) speculate whether microfinance institutions would be better able to serve the poor if the industry were to become more commercialized. The reasoning behind this suggestion is that the profit motives lead them to become more efficient and willing to explore new markets for their products. Furthermore, her study finds that microfinance institutions with good performance were able to tailor their delivery methods to the poor so efficiently that the clients could afford the full cost of repayment, ultimately leaving the institutions to fulfil both missions of outreach and sustainability. Littlefield, Morduch and Hashemi (2003) support this view in their findings that performance is better in microfinance institutions where poorer customers are targeted. This study does lack generality, but nevertheless contradicts the claims of microfinance mission drift.

Mersland & Strøm (2013) adds to the debate on microfinance mission drift by stating that the level on the interest rate given to microfinance customers and the profitability obtained by the institutions must be connected to the costs of providing the loans. Rhyne (1998) find that there is a connection between the financial viability of the microfinance institution and their willingness to set interest rates at levels that fully recover costs, and claims that those who do not set interest rates at such a level thus chose to remain dependent on subsidy. One interpretation of this is that some microfinance institutions are consequently subsidizing interest rates to their clients. The effect of the subsidy on interest rates and outreach to customers are too comprehensive to include in this thesis, but we do recognize that whether or not the microfinance institution is receiving subsidy may lead them to increase outreach on the expense of financial sustainability.

Although Mersland & Strøm (2010) find no mission drift when viewing the microfinance industry as a whole, they do find that the size of average loans increase with increased average profits and average costs, suggesting that mission drift tendencies may be neutralized by a cost-efficiency focus in microfinance institutions. This builds to Rhyne's (1998) findings that the social- and financial objectives of the institutions are in fact complementary, and above all that financial sustainability serves their mission of outreach to the poor. It becomes evident that the industry would take advantage from improving their cost efficiency. This thesis will investigate whether institutions can encounter this need by adjusting their tolerance for loan defaults.

3.3 Conceptual framework and research hypothesis

Based on evidence from past research and the theoretical background this chapter will present the hypothesis and conceptual framework of the thesis. The purpose of the conceptual framework is to illustrate the hypothesized relationship between the default on loans rate and the operating cost rate in microfinance institutions.

Past research shows that the relationship between default on loans and operating costs run in both directions (see Berger & de Young, 1997), meaning that default on loans can influence the operating costs rate, but the operating costs rate can also influence the level default on loans. This thesis is interested in the nature of this relationship, that is, the functional form of it. The hypothesis is based on the idea that the microfinance institutions today may be too concerned with maintaining low levels of default, perhaps to the degree to which they would be better off by easing up their strong focus on this and allow for their resources to be spent differently. The expectation is that there is a non-linear relationship between the two variables, and that this might be quadric and convex. That means that we expect that very low levels of default will lead to higher operating costs, due to great resources being spent to hold such levels. We expect that the cost will decrease as default rates increase because resources are freed. Yet, at some point the costs of defaults will have to exceed the savings gained by lightening the use of resources, so we expect to see that further default rates will lead to increasing operating costs again. The expected relationship can be illustrated as follows:

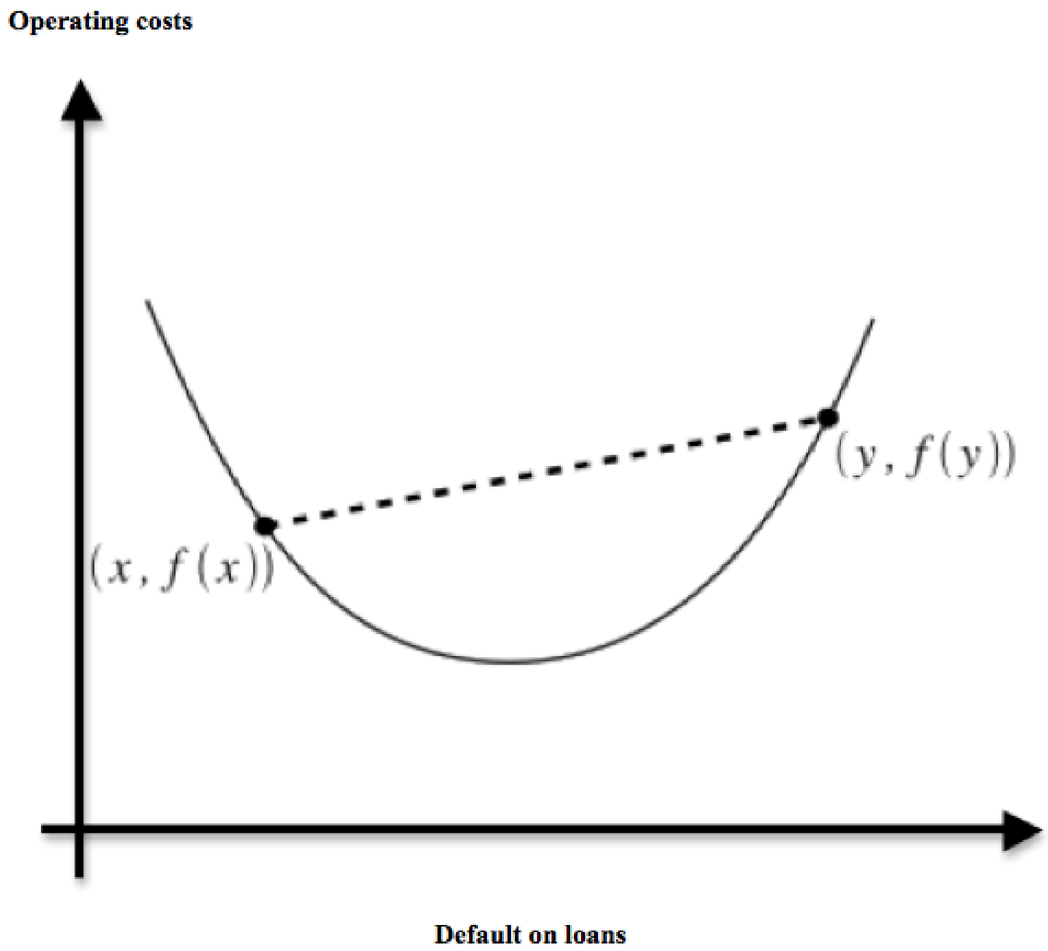


Figure 3.1 Expected relationship between the dependent and independent variable

Thus, our hypotheses can be formulated:

H_1: There exist a statistically significant relationship between default on loans and operating costs in microfinance institutions.

H_{A1}: There does not exist a statistically significant relationship between default on loans and operating costs in microfinance institutions.

H_{2a}: There exist a non-linear relationship between default on loans and operating costs in microfinance institutions.

H_{A2a}: There does not exist a non-linear relationship between default on loans and operating costs in microfinance institutions.

H_{2b}: There exists a quadric, convex relationship between the default on loans rate and the operating costs rate in microfinance institutions.

H_{A2b}: There does not exist a quadric, convex relationship between the default on loans rate and operating costs rate in microfinance institutions.

The relationship is expected to be influence by several other factors, which is portrayed in the conceptual framework below and elaborated on in chapter 4.2.3.

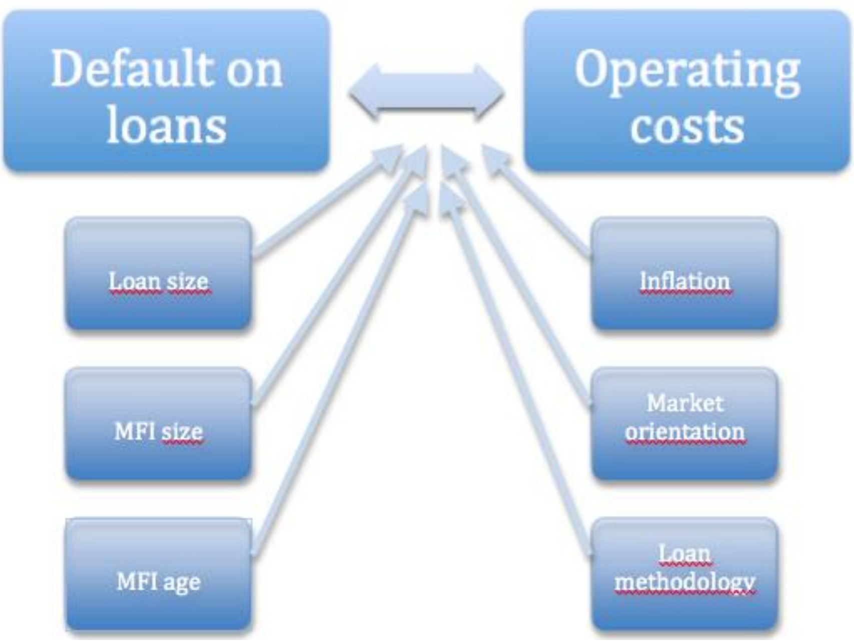


Figure 3.2 Conceptual framework

4. Data

4.1 Data and Sample

This study is based on data collected by Mersland, referred to as the Mersland data base. The dataset covers financial and general data on 463 microfinance institutions in 77 countries collected from risk evaluation reports by specialized rating agencies supported by the Rating Fund of Consutative Group to Assist the Poor: MicroRate, Microfinanza, Planet Rating, Crisil and M-Cril. The data set resembles Garmaise and Natvidad (2010), but contains over three times as many microfinance institutions in above twice the number of countries. The rating reports can be found at www.ratingfound2.org. Each rating was performed during the period 1996-2012. The data and its quality is well recognized in the academics and has been used as the basis for several academic articles published in development, management, economic and financial journals (see e.g. Mersland & Strøm, 2008; Mersland, 2009; Hartarska, Shen & Mersland, 2013; Mersland, & Urgeghe, 2013). An overview of the institutions and the countries represented in the data set is offered below.

Overview of countries and number of microfinance institutions

Country code	Country	No. of Microfinance institutions	Country code	Country	No. of Microfinance institutions
1	Albania	3	**39**	Russian Federation	15
2	Argentina	1	**40**	Senegal	11
3	Armenia	2	**41**	South Africa	3
4	Benin	8	**42**	Sri Lanka	2
5	Bolivia	6	**43**	Tanzania	8
6	Bosnia Hercegovina	10	**44**	Togo	4
7	Brazil	14	**45**	Trinidad and Tobago	1
8	Bulgaria	2	**46**	Tunisia	1
9	Burkina Faso	3	**47**	Uganda	14
10	Cambodia	14	**48**	Montenegro	2

11	Chile	2	**49**	Cameroun	5
12	Colombia	12	**50**	Guinee	1
13	Dominican Republic	5	**51**	East Timor	1
14	Ecuador	18	**52**	Bangladesh	2
15	Egypt	5	**53**	Nepal	5
16	El Salvador	6	**54**	Vietnam	3
17	Ethiopia	10	**55**	Azerbaijan	8
18	Georgia	7	**56**	Mongolia	3
19	Guatemela	8	**57**	Nigeria	5
20	Haiti	3	**58**	Mozambique	1
21	Honduras	10	**59**	Tajikistan	9
22	India	32	**60**	Croatia	1
23	Indonesia	4	**61**	Chad	1
24	Jordan	3	**62**	Rwanda	5
25	Kazakhstan	4	**63**	Zambia	3
26	Kenya	12	**64**	China	4
27	Kyrgyzstan	5	**65**	Serbia	1
28	Madagascar	3	**66**	Ghana	5
29	Mali	5	**67**	Malawi	1
30	Mexico	21	**68**	Gambia	1
31	Moldova	2	**69**	Kosovo	4
32	Morocco	7	**70**	Rep of CongoBrazz	1
33	Nicaragua	14	**71**	Burundi	1
34	Pakistan	1	**72**	Niger	5
35	Paraguay	2	**73**	DRC - Kinshasa	1
36	Peru	39	**74**	Afghanistan	1
37	Philippines	15	**75**	Costa Rica	1
38	Romania	2	**76**	Lebanon	2
			77	Turkey	1
Total number of microfinance institutions:					463

Table 4.1 Overview of countries and microfinance institutions

Despite covering many microfinance institutions in several countries, the data set is not a flawless representative of the microfinance industry. The representation of very large microfinance institutions is relatively limited and it does not succeed in fully covering the next to infinite number of small savings and credit cooperatives. Nonetheless, the characteristics of the microfinance institutions in the Mersland dataset are found to be fairly similar to other publically available data, such as the larger MIX Market (www.mixmarket.org) (Mersland & Strøm, 2012b), and it has the benefit of being gathered by third parties.

Below is an overview of the main characteristics of the microfinance institutions in the dataset. The overview presents the mean, minimum and maximum rates/amounts reported, as well as the standard deviation, median and number of observations. All outliers have been trimmed away so that the averages represent the mainstream microfinance institutions.

	Mean	**Median**	**Std.**	**Min.**	**Max.**	**Count**
Average loan ($)	1,204	743	1,573	15	18,984	1,322
Credit Clients	17,971	7,028	33,710	10	394,462	2,191
Assets ($)	11,268,680	3,766,000	24,744,468	0	279,350,811	2,291
Loan Portfolio ($)	8,183,922	2,771,352	17,070,439	3,425	156,789,000	2,298
Equity Fraction	39.55%	34.04%	25.14%	0%	100%	2,220
Portfolio Yield	37.79%	34.10%	18.45%	0.7%	127.7%	2,194
Operating cost of portfolio	30.55%	22.40%	26.95%	2.93%	351.0%	2179
Par30	5.56%	3.20%	7.51%	0.00%	54.70%	1.981
Portf. write-off	2.79%	1.40%	4.36%	0.01%	42.0%	2,060
Return on assets	1.36%	2.95%	12.04%	-99.00%	34.20%	2,186
Adjusted ROA	-1.66%	-0.20%	8.73%	-43.60%	22.30%	1,102

Table 4.2 Characteristics of microfinance institutions

The overview reveals that the average loan disbursed by the microfinance institutions in our data set is approximately U.S. $1,204. Moreover, we notice that the minimum average loan disbursed by a microfinance institution is as low as U.S. $15. This illustrates that the average loan size is very small compared to what is common in commercial banks. The median of the reported numbers (U.S. $743) is even lower that the mean, which tells us that even though there are some “mega” microfinance institutions, the mainstream stay true to the original intent of microfinance; offering small loans to the poor.

When it comes to the number of credit clients, the mean in the dataset is approximately 17,971 clients per microfinance institution in any year. The median is however much smaller and tells us that the typical microfinance institution is small in terms of number of credit clients, although exceptions are present.
As will be discussed in chapter 4.2.3 the total assets are one of the most common measures of the microfinance institutions' size. The table above show that the institutions in our data set average around U.S. $11,3 millions over the years reported, and that the variation between the cases is great; from zero at the lower end to over U.S. $279 million on the upper end. The median does on the other hand tell us that the main part of the cases reported are of a smaller size, around U.S $3,76 million.

The loan portfolios in our dataset averages around U.S. $8,18 million, but is spread out from only U.S. $3,425 on the lower end to over U.S $156 million on the upper end. The median only about 1/4 of the mean, which tells us that, as was true for the previous characteristics, the mainstream of the cases is in the lower end.

The equity fraction displays the level of equity related to total assets, as reported on each microfinance institution's balance sheet. The loan portfolio makes up the greatest part of the total assets. We see that both the mean and the median shows satisfying levels, which tells us that the microfinance institutions in our data set in general have been well capitalized over the measured periods.

The table shows that in our dataset, the portfolio yield, which proxies the lending rate is on average 37.79%. This characteristic is mainly interesting viewed together with the operating cost per portfolio, which on average is 30.55%. We notice that the operating costs of the portfolio is not far from the same level as the portfolio yield, which supports the arguments presented in chapter 2.3, that the lending rates to a great degree is made up of high operating costs.

We also notice that the portfolio at risk (30 days) is 5.95% on average, which is also eating up parts of the portfolio yield. Consequently, the return on assets is low, averaging at 1.36% with a somewhat better median of 2.95%. The adjusted return on assets is low as well, averaging at -1.66% with a median of -0.20%. We notice that both for return on assets (ROA) and adjusted return on assets (AROA) the median is slightly above the mean, which tells us that the mainstream of microfinance institutions experience a somewhat better return that what is illustrated by the mean.

4.2 Variables

The purpose of this study is to determine whether there is a relationship between the default on loans rate and the operating costs rate in microfinance institutions and what the functional form of this relationship looks like. To ensure that this is estimated in the best way possible, several measures of both the default on loans-concept and the operating costs-concept was evaluated before the measures best suited for this study were chosen. A number of control variables that are assumed to have an effect on either default on loans rate or the operational cost rate will be included. The choice of variables is partly based on previous studies and partially based on own judgements.

4.2.1 Dependent variable

The operating expenses related to assets will be used as the measure for evaluation of the microfinance institutions' the operating costs rate, as it is a well-known and commonly accepted measure for that purpose. The measure states the ratio of operating expenses to annual average total assets, where annualized figures are used if the report gives figures from within a year, using the formula:

$$\frac{\text{Operating expenses}}{\text{Annual average total assets}}$$

The annual average total assets is calculated as follows:

$$\frac{\text{Total loan portfolio year X + Total loan portfolio year (X-1)}}{2}$$

Mersland & Strøm (2013) find that as the operating costs per client decreases, the microfinance institutions are able to grant lower average loans, and thus reach out to more people. This serves as a motivation for keeping operational costs as low as possible, as Mersland & Strøm (2010) find that microfinance institutions with the best potential to reach poor customers are also the ones that are most efficient.
A basic assumption for the use of multivariate analysis is that the shape of the data distribution for an individual metric variable corresponds to a normal distribution (Hair, Black, Babin & Anderson, 2010). Therefore, for all the variables a histogram was generated, displaying the normality line. For the variables not displaying normal distribution, a transformation was carried out using gladder tests to see if any of the suggested transformations provided a satisfying distribution. The results of the distribution tests pre- and post transformation can be found in appendix 1. The tests showed that the logarithmic function of operating expenses related to assets is best suited for our model, presented as: *lnoperexp_assets* in data outputs.

In addition to using operating expenses related to assets, this thesis will include an alternative measure for the operating costs rate, namely the logarithmic function of operating expenses related to portfolio, *lnoperexp_portf.* The purpose of including this additional measure is to check the robustness of the findings in terms of whether they can be supported by the use of other measures. If the measures provide similar results, the strength of the research will improve and the probability of drawing good conclusions will advance.

The measure states the ratio of the operating expenses to the annual average loan portfolio, where figures are annualized if the report gives them from within a year, using the formula:

$$\frac{\text{Operating expenses}}{\text{Annual average total loan portfolio}}$$

The average total loan portfolio is calculated as follows:

$$\frac{\text{Total loan portfolio year X + Total loan portfolio year (X-1)}}{2}$$

Results of regression using this alternative dependent variable can be found in appendix 2.

4.2.2 Independent variable

When estimating the microfinance institutions' operating costs it is important to incorporate risk, which is commonly measured using non-performing ratios such as the portfolio at risk. Portfolio at risk (Par30) is an uncertainty measure, displaying the ratio of loans that are 30 days or more in arrears (Mersland & Strøm, 2012a). This consideration is necessary because high levels of non-performing loans demand added resources to administer the risk. The loan portfolio quality is very important to the performance of the microfinance institutions, as it represents one of their largest assets. Because the loans in general are not backed with bankable collateral, such as a mortgage on a house, etc., the risk associated with poor management of the loan portfolio can be very dramatic (Jansson, 2003).

Hartarska et. al., 2013 find that costs increase with risk independently of whether output is measured in dollars or in number of active clients, but Albuntas, Carbo, Gardener & Molyneux (2007), on the other hand, argue that risk is inversely related to inefficiency and that in most cases cost inefficiency is positively related to asset size. Mersland & Strøm (2013) further argue that some microfinance institutions may use vast resources in effort to obtaining repayment on their loans, and that higher default rates in these cases may reduce operating costs.

Based on existing theory we expect to see a relationship between default on loans and operating costs in microfinance institutions, and that this relationship is non-linear. Our study uses the logarithmic function of Par30, presented as *lnpar30* in the data output.

In addition to using Par30, this thesis will include an alternative measure for the default on loans rate, namely the logarithmic function of the combined credit risk, *lncomb_credrisk.* This measure is computed combining the write-off ratio and the Par30 in each microfinance institution. The purpose of including this additional measure is to check the robustness of the findings in terms of whether they can be supported by the use of other measures. If the measures provide similar results, the strength of the research will improve and the probability of drawing good conclusions will advance.

4.2.3 Reversed causality

Microfinance institutions that are less efficient may be tempted to take on greater levels of risk to counterweigh lost returns (Albuntas et al., 2007). On the other hand, efficiency can be influenced by the level of risk the microfinance institution undertakes. Berger & De Young (1997) find that commercial banks with high levels of non-performing loans also experience declining cost efficiency. This is consistent with the theory that extra monitoring and administration of risky portfolios lead to higher operating costs. So far, to my knowledge, no study on this relationship has been conducted in the microfinance industry. As argued in chapter 2, we do however expect to find similar results for microfinance institutions. Berger & De Young's (1997) data further insinuate that low levels of cost efficiency lead to increases in nonperforming loans, consistent with the theory that inefficient firms may appeal to undertaking more risk. Because theory tells us that there might be issues concerning reversed causality/endogeneity, instrument variables will be included in the model. Endogenity is a well-known and persistent problem in research on corporate governance (Bøhren & Strøm, 2007) that makes it difficult to distinguish cause from effect. This phenomenon relates to the independent variable explaining the dependent variable, while the dependent variable in turn explains the independent variable (Dahlum, 2014).

The problems concerning endogeneity can be overcome using predetermined variables (Kang & Sivaramakrishnan, 1995), such as the first and second lag of the independent variable, as instrument variables. These variables are found in the IV regression in chapter 6.3 as *lagPar30* and *lag2Par30*, respectively.

4.2.3 Control variables

Loan size

Because microfinance institutions serve a social mission of outreach to poor people, small loans of short duration is common in the industry. With the smaller loans banks are able to extend credit to more people, including those who are only able to repay very small amounts, i.e. the poor people. The size of loans is thus a commonly used measure of microfinance institutions' outreach and fulfilment of their social mission (Bhatt & Tang, 2001; Cull, Demigüc-Kunt & Morduch, 2007). In this study the size of loans will be included as a control variable as it is thought to affect the operating costs as well as on the microfinance institution's risk in several ways. Firstly, the banks incur a fixed element in loan provision expenses, making smaller loans relatively more costly compared to larger loans. Secondly, given a fixed amount of resources, outreach to more of the poorer customers will compromise larger business with the more fortunate customers, leading to extra costs in terms of risk assessment of new customers and compensation for lost business (Mersland & Strøm, 2012a). Thirdly, the smaller loans provide the microfinance institution with a way of risk diversification, as credit is spread out on a large amount of borrowers.

The size of loans, or the loan amount, will be represented by the average disbursed loan amount, namely the variable *lnloan_disb_av.* The average loan is defined as the loan portfolio divided by the number of credit clients in the institution, and is thus a usage measure. Based on existing theory our prediction is that we will see a negative effect on the operating costs of the microfinance institutions when the average size of loans increases.

Microfinance institution size

The size of the microfinance institution will be included as a control variable as it is reasonable to assume that larger institutions will accumulate scale advantages that enable them to obtain lower operating costs. This assumption is supported in Mersland & Strøm (2013) where they find that the size of the microfinance institution, measured in total assets, will have a negative effect on the operating costs rate of that institution. Total assets is the main and commonly accepted measure of company size used in finance, and is available and well suited in this study. We will correct for country-specific traditions and influence by adjusting total assets to the GDP per capita in each country, as suggested by Aguilera and Jackson (2003). The variable is named *lnTotalassets_GDPadj* in the data output.

Allen and Rai (1996) find clear scale advantages in financial institutions. Because of the tremendous growth in the microfinance sector, the institutions are expected to achieve cost savings in the future, resulting in further lending by means of small loans and ultimately a greater outreach to the poor. Moreover, with increasing competition the banks that have achieved scale will have a competitive advantage, which further supports an expectation of a negative relationship between microfinance institution size and operating costs. In addition to scale advantages larger microfinance institutions are thought to have an advantage in popularity and reputation, which in turn will attract new customers without much effort from the microfinance institution. Thus, based on existing theory we expect to see a negative relationship between the size of the microfinance institution and its operating costs rate.

Microfinance institution age

Mersland & Strøm (2010) propose that the microfinance institutions gain experience through its daily operations, recurring interactions with clients and market transactions, and that they over time will accumulate cost-effective ways to run their business. Thus, we can expect that the microfinance institutions will reduce its operating cost rate over as they age, and that the regression will show a negative relationship between the age and the operating cost rate. This expectation is further supported by Caudill et al. (2009) who find that the cost efficiency in microfinance institutions increase over time.

Because of this, the MFI age is included as a control variable and named *sqrt_age* in the data output. The microfinance institution age is derived by subtracting the original start-up of the organization from the year. The term is then squared to meet the assumptions of multivariate regression.

Inflation

The study controls for regional factors by including the countries' inflation rate. Some regions may have suffered particularly severe economic upturns or downturns relative to the rest during parts of the sample period. This is expected to influence the operating costs of microfinance institutions located in those regions. Moreover, including the inflation rate and benchmarking the MFI size against the GDP per capita will help control for country heterogeneity.

The study uses the logarithmic function of the inflation, marked *lninflation*, and states the inflation in the country at the end of a given period. Because the operating costs in any microfinance institution is contingent on the general price level of the country, we expect to se a positive relationship between the countries' inflation rate and the operating cost rate in the corresponding MFIs.

Market orientation

United nations (2006) reveals that the rural areas are where poverty is most concentrated. Reaching the rural areas should thus be a significant goal in microfinance, in accordance with the social mission of outreach. People living in rural areas are generally in more financial need than those in urban areas. Rural areas are thus harder for microfinance institutions to enter and operate in. It can be argued that it is costly to serve rural clients because they lack skills regarding microenterprises; hence, microfinance institutions may offer literacy/training services, resulting in higher costs. Mersland & Strøm (2013) find that higher operating cost per client leads microfinance institutions to seek customers in urban communities. Consequently, their study suggests that there is a positive relationship between outreach to rural areas and operating costs.

Based on these arguments we expect to see a positive relationship between the outreach to rural areas and operating costs, i.e. that microfinance institutions with outreach to rural areas incur higher operating costs. In our model the market orientation is included using the dummy variable *rural*. This variable has the value "1" for microfinance institutions that serve rural customers and "0" for those who do not.

Loan methodology

As a response to the lack of collateral and reduced chances of legally enforcing repayment, the microfinance industry commonly issue group loans where the loans are given to individuals, but whole groups are responsible for the repayment of it (Armendàriz and Morduch, 2010). The idea is that the social capital implied by being part of a group substitutes collateral (Tirole, 2006), and that group members will monitor other members (Varian, 1990). This would lower costs for microfinance institutions in terms of monitoring and control. However, Mersland and Strøm (2010a) find that the benefit of group lending is often outperformed by the cost associated with it.

The loan methodology is included as a control variable in the model using the dummy *DM_individ*, that states whether or not the microfinance institution offer individual loans; the value "1" being "yes" and "0" being "no". Existing theory is somewhat inconclusive about the relationship between individual lending and operating cost, and we can expect the relationship to be either negative or positive. A negative relationship would imply that microfinance institutions that offer loans to individuals incur lower operating costs, whereas a positive relationship would imply that operating costs increase with loans offered to individuals.

Table 4.3 below provides an overview of the variables used in this study.

Variable	**Definition**
Dependent variables	
ln Operating expense/assets	The natural logarithm of operating expenses related to annual average total assets
ln Operating expense/portf	The natural logarithm of operating expenses related to annual average loan portfolio
Independent variables	
ln PaR30	Portfolio at risk (30days). States the ratio of loans that are 30 days or more in arrears.
Control variables	
Loan size	The natural logarithm of average disbursed loan amount. The average loan is defined as the loan portfolio divided by the number of credit clients.
MFI size	The natural logarithm of total assets adjusted for country GDP per person.
MFI age	The squared function of original start-up of the organization (establ_activ) subtracted from year (yr).

Inflation	The natural logarithm of the inflation in the country at the end of a given period as indicated in the report.
Market orientation	Measures outreach to rural areas. The value “1” for institutions that serve rural customers and “0” for those who do not.
Loan methodology	States whether or not the microfinance institution offer individual loans; the value “1” being “yes” and “0” being “no”.

Table 4.3 Overview of variables

5. Method

Research is a process that aims at providing new knowledge and increased understanding through systematic efforts, and is used by businesses to ensure clever, well-versed decisions (Joyner, 2013). This chapter will present the research design, analytical method and variables used in my research on the relationship between loan default and operating costs in microfinance institutions.

5.1 Research design

The research design provides a framework that identifies the methods and procedures for collecting and analysing the data that will be used in the research, and offer an outline of actions to be made (Joyner, 2013). Joyner (2013) hold that objectives of the study identified in the first stages of the research should be incorporated to the research design to make certain that the data used is suitable for the particular research problem. This study aims to portray the relationship between variables and to say something about cause and effect to these variables. More specifically the study investigates the effect on the operating cost rate caused by default rates in microfinance institutions. Causal research attempts to establish the effect of an action (Joyner, 2013) and is thus the appropriate research design for this study.

5.2 Regression analysis

As regression analysis is a very versatile dependence technique, it is the most commonly used and is applicable in all aspect of business decision-making (Hair, Money, Samouel & Page, 2007). It is used to measure the linear dependency between a dependent variable and one or more independent variables, and is widely accepted as an appropriate tool in identifying causal relationships (Joyner, 2013)

5.2.1 Multiple regression analysis

Multiple regression analysis is a general statistical technique used to analyse the relationship between a dependent variable and a number of independent variables. It can be formulated as:

$Y_1=\beta_0 +\beta_1X_1+\beta_2X_2+....+\beta nXn+\varepsilon$

where all variables are metric.

Y represents the dependent variable and X the different independent and control variables. β_0 is the intercept, while β_1, $\beta_2 \ldots \beta_n$ represents the slope. The beta coefficients, β_n are standardised coefficients that permit comparison between coefficients as to their relative explanatory power of the dependent variable (Hair et al., 2007). The set of these weighted independent variables structure the regression variates, also referred to as the "regression model" or "regression equation", which is a linear organization of the independent variables that most precisely predict the dependent variable (Hair et al., 2007). The error term, ε, represents all other unobservable factors that may affect Y, so called "noise". This is the degree to which the data values do not truly measure the characteristics of the dependent variable (Hair et al., 2007).

5.3 Ordinary Least Squares (OLS)

The OLS is a mathematical technique used to ensure that the regression line used to identify the relationship between the dependent variable and the independent variable produce the smallest possible total error. More specifically the model generates a straight line that minimizes the sum of squared deviations of the actual values from its predicted regression line (Joyner, 2013). The deviations are squared to control for positive and negative faults cancelling each other out. In the OLS model the deviations of observations from the regression line are represented by the symbol *e*, and no other line can produce less error. The OLS criterion is as follows (Joyner, 2013):

$\sum^{n}_{i=1} e_i^2$ is minimum where:

$e_i = Y_i - \hat{Y}i$ (the residual)

Y_i = actual observed value of the dependent variable

$\hat{Y}_i$ = estimated value of the dependent variable

N = number of observations

i = number of the particular observation

5.4 Prerequisites for multiple regressions

As each of the independent variables has some explanatory power on Y, more variables will always cause a larger portion of Y to be explained. Enhancements in prediction of the dependent variable can thus be made by adding more independent variables and even transforming them to denote aspects of the model that are not originally linear (Hair et al., 2007). To do so, several assumptions about the relationship between the variables must be made (Chen, Ender, Mitchell & Wells 2003):

1. Normality of the error term distribution
2. Collinearity
3. Independence of error terms (autocorrelation)
4. Constant variance of error terms (homoscedasticity)
5. Linearity of the phenomenon measured

These assumptions will be further explained below and tested in chapter 6.1.

5.4.1 Normality

The assumption of normality refers to the shape of the data distribution for an individual metric variable and its equivalence to the normal distribution. If the deviation from the normal distribution is considerable, all consequential statistical tests are invalid as normality is necessary to do the t and F statistics (Hair et al., 2010). The normality of a variable distribution can be checked for by generating a histogram with a normality plot and by analysing the kurtosis and skewness values of each variable.

Kurtosis refers to the level of peak or flatness of the distribution compared to that of the normal distribution. If the distribution is more peaked than the normal distribution, it is said to be leptokurtic (Hair et al., 2010). In this case the value of the kurtosis will be greater than 3.00 (Acock, 2012) and the tails of the distribution will be too thin. When the distribution is flatter than the normal distribution and the tails are too thick, we will see a kurtosis value of less than 3.00 (Acock, 2012) and we say that the distribution is platykurtic (Hair et al., 2007). In multivariate data analysis kurtosis values up to 10.0 is considered acceptable when evaluating the normality of a variable (Acock, 2012).

5.4.2 Collinearity

Collinearity is concerned with the linear relationship between variables. When more than two variables are involved it is referred to as multicollinearity. The main concern with multicollinearity is that as the degree of it increases, the regression model estimates of the coefficients become volatile and the standard errors for the coefficients can get uncontrollably inflated (Chen et al., 2003). Increases in multicollinearity reduce the overall R^2 that can be achieved, cofound estimation of the regression coefficients and negatively affect the statistical significance tests of coefficients (Hair et al., 2010).

5.4.3 Independence of error terms (autocorrelation)

The independence of error terms is about making sure that errors linked with one observation are not correlated with the errors of any other observation (Chen et al., 2003). This assumption is important in time series studies where it is probable that errors for observations between contiguous periods will be more highly correlated than for observations more separated in time. This phenomenon is referred to as autocorrelation.

5.4.4 Constant variance of error terms (homoscedasticity)

Homogeneity of variance of the residuals is among the most important assumptions for the ordinary least squares regression. Homoscedasticity refers to the assumption that dependent variables demonstrate equal levels of variance transversely in the assortment of predictor variables (Hair et al., 2010) and is advantageous because the variance of the dependent variable explained in the dependence relationship should not be restricted to a limited range of the independent values.

5.4.5 Linearity of the phenomenon measured

When conducting linear regression, the relationship between the response variable and the predictors should be linear. If this assumption is not withheld, the linear regression will try to fit the model into a straight line anyways. It is therefore necessary to control for linearity and make adjustments to the regression model if the relationship between the dependent and independent variable is non-linear. Tests and adjustments are carried out in chapter 6.1.

5.5 Panel Data

Longitudinal data or Panel data is defined as continual measurements on the same individual unit at different points in time, making it possible to detect variation over time as well as variation over unit (Cameron & Trivedi, 2010). Cameron & Trivedi (2010) further explains that panel data can be either balanced, meaning that all individual units are observed in all time periods ($T_i = T$ for all i), or unbalanced. This study is based on yearly reports from 463 Microfinance institutions in the period between 1996-2012, where the range of institutions has been consistent over the years, thus making the data set suitable for an analysis using panel data. The advantage by using the same measurement parameters at the different points in time is that it allows us to detect relationships and variations in the sample that would otherwise not be apparent. It further permits the researcher to control for variables that cannot be observed or measured, like differences in business practices across companies or cultural factors (Torres-Reyna, 2007). Cameron & Trivedi (2010) emphasize on two main methods for analysing panel data, namely “fixed effects” and “random effects”.

5.5.1 Fixed effects

The fixed effects method investigates the relationship between the independent, explanatory variable and the dependent variable in a single unit. Each unit is defined with individual characteristics that may affect the explanatory variables and may impact or bias the findings the analysis provides. This impact or bias rationalizes the assumption of the correlation between the unit’s error term and explanatory variable (Torres-Reyna, 2007). The model removes the effect of time-invariant characteristics and allows us to assess the net effect of the independent variable on the dependent variable, and is thus suited when one is concerned with merely analysing the impact of variables that vary over time. As each unit is different, its error term and individual characteristics, captured by its constant, should not be correlated with other units (Torres-Reyna, 2007). (Torres-Reyna, 2007) expresses the model of fixed effects:

$$Y_{it} = \alpha_I + \beta_1 X_{it} + u_{it}$$

Where:

α_I (i=1….n) is the unknown intercept for each unit (n = unit-specific intercepts)

Y_{it} is the dependent variable where i = unit and t = time.

X_{it} represent one independent variable

β_1 is the coefficient for the independent variable

u_{it} is the error term

The intercept, α_I, captures the effect of the individual characteristics of the units and will be constant over time. The independent variable, X_{it}, is weighted by β_1, which measures the effect of the independent variable on the dependent variable, Y_{it}.

This model is limited by the degrees of freedom, which tends to be low. This is because one degree of freedom is lost per cross-sectional observation by eliminating the properties that are time-invariant. Additionally, the explanatory variables that do not vary over time in each unit have a perfect collinearity with fixed effects, making them and their coefficients unfit for the model (Joyner, 2013)

5.5.2 Random effects

The random effects method assumes that the variations across units are random and uncorrelated with the explanatory or independent variables in the model (Torres-Reyna, 2007). Greene (2012) explains: "the crucial distinction between fixed and random effects is whether the unobserved individual effect embodies elements that are correlated with the repressors in the model, not whether these effects are stochastic or not". Each cross sectional unit has an intercept given through a distribution that is centred in an average intercept (Studenmund, 2011), leaving the intercept independent of the error term for each individual observation. Torres-Reyna (2007) expresses the model of random effects:

$Y_{it} = \alpha_I + \beta_1 X_{it} + u_{it} + \varepsilon_{it}$

The intercept, α_I, in this model is based on the normal distribution of all the units. Like in the fixed effects model the independent variable, X_{it}, is weighted by β_1, which measures the effect of the independent variable on the dependent variable, Y_{it}. u_{it} represents the between-unit error, whereas ε_{it} represents the within-unit error. One assumption of the random effects is that the unit's error term is uncorrelated with the predictors, thus allowing for time-invariant variables to act as explanatory variables (Torres-Reyna, 2007) and it is necessary to describe the individual characteristics that may influence the predictor variables. Problems may occur if some variables are unavailable, leading to variable bias being left out of the model (Studenmund, 2011).

Nevertheless, here are some evident benefits of selecting the random effects method over the fixed effects method. Firstly, the random effects model will have a higher degree of freedom as it estimates parameters that describes the distribution of the intercepts rather than estimating the intercept for each unit (Studenmund, 2011). Secondly, the model allows you to estimate the coefficients of the predictors that are constant over time.

5.5.3 Hausman test

The Hausman test can be used to investigate whether the fixed effects model is appropriate for analysing a dataset by investigating how the independent variable and the intercept are correlated (Hausman, 1978). The principle behind the test is to compare two predictors where one is fixed both in the null hypothesis and in the alternative hypothesis, and the other fixed only in the null hypothesis (Verbeek, 2012). The null hypothesis is that the random effects model is the preferred method, whereas the alternative hypothesis is that the fixed effects fixed effects model is preferred (Greene, 2012). The test identifies whether the individual error terms correlate with the independent variables. If they do not correlate the random effects model is the most appropriate method. The decision is based on an evaluation of the p-value, which is defined as prob > chi2 in the results table of the Hausman test (Torres-Reyna, 2007).

If the measured p-value is greater than the selected significant level, the random effects model will be most appropriate to the tested dataset. To control for fixed effects, a Hausman test will be conducted on the basis of a model with fixed effects and a model with random effects. Based on the results, one of the models will be chosen.

6. Results and analysis

The ability of an added independent variable to improve the prediction of the dependent variable is linked to the correlations to the additional independent variables already in the regression equation, as well as to its correlation to the dependent variable (Hair et al., 2010). Collinearity expresses this relationship and is measured as the correlation between two independent variables. This implies that two variables are near perfect linear combinations of one another (Chen et al., 2003). When the correlation concerns three or more independent variables (evidenced when one variable is regressed against the others) it is called multicollinearity. Hair et al. (2010) affirm that: “the impact of multicollinearity is to reduce any single independent variable’s predictive power by the extent to which it is associated with the other independent variables”. When the relationship among the predictors is (near) perfectly linear, the estimates for a regression model cannot be distinctly calculated.

To control for this, a correlation analysis will be executed. Hair et al. (2010) explains that bivariate correlations of 0.70 or higher generally is considered to be problematic. Additionally, even lower correlations can be problematic if they are greater than the correlations between the independent and dependent variables. Nevertheless, it is commonly accepted to allow values up to 0.90 (see e.g. Tabachnick & Fidell, 2001; Tamhane & Dunlop, 2000) in the correlation matrix as further tests of multicollinerarity will confirm/expose whether this is an issue in the model. Below is an illustration of the correlation analysis.

	lnoper~s	lnpar30	lnloan~v	lnTota~j	sqrt_age	lninfl~n	rural	DM_ind~d
lnoperexp_~s	1.0000							
lnpar30	-0.0573	1.0000						
lnloan_dis~v	-0.4150	0.0065	1.0000					
lnTotalass~j	-0.2702	-0.1236	0.0221	1.0000				
sqrt_age	-0.0998	0.3825	-0.0247	0.0377	1.0000			
lninflation	0.1165	0.0806	-0.0517	-0.0153	0.0249	1.0000		
rural	-0.2586	-0.0495	-0.0797	-0.0645	-0.0530	-0.0529	1.0000	
DM_individ	-0.1305	0.0195	0.3244	0.0433	-0.0373	-0.0367	-0.1683	1.0000

Table 6.1 Correlation matrix

Table 6.1 above show that all correlations are well below the generally accepted limit of 0.70 and will remain in the model for further analysis.

6.1 Test of assumptions underlying the regression analysis

Before the regression is run it is practical to test the underlying assumptions of linear regression, namely: normality, collinearity, homoscedasticity, independence of error terms and linearity. The results of the tests will be displayed below with comments about required corrections and adjustments.

6.1.1 Normality

To ensure that the p-values of the coefficient can be considered reliable, the error terms of the variables have to be normally distributed. This can be controlled for by predicting the residuals, r, before running a Kennel density estimate of these. The Kennel density estimate will compare the distribution of the residuals to the normal distribution. Additionally, a standardized normal probability plot, *pnorm*, which displays the distribution function, and quantiles plot, *qnorm*, which gives the quantile function, will be included to confirm or discredit the results of the Kernel density estimate. The *pnorm* provides the cumulative probability distribution at a specified value of x, and will be interpreted by looking at deviations from the straight line that represents the normal probability. The *qnorm* graphs the quantiles of a variable against the quantiles of a normal distribution, and like the *pnorm* it is interpreted by looking at the deviations from the straight line representing the normal distribution.

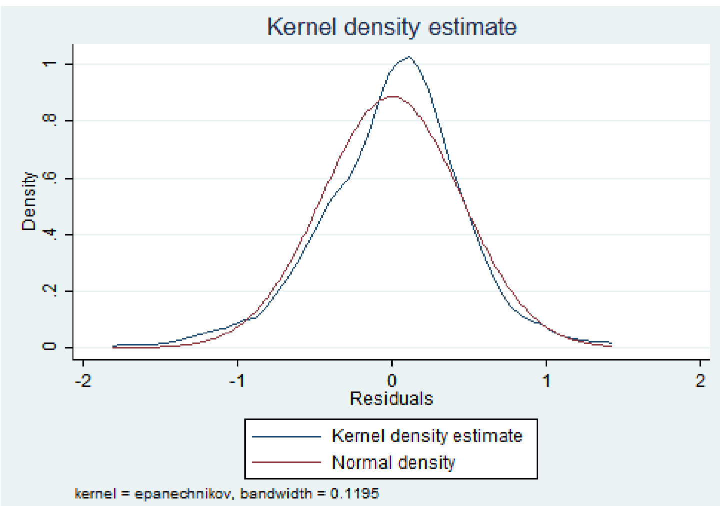


Figure 6.1 Test of normality

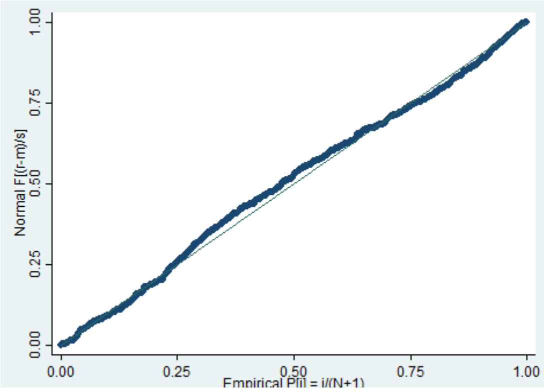


Figure 6.2 Standardized normal probability plot

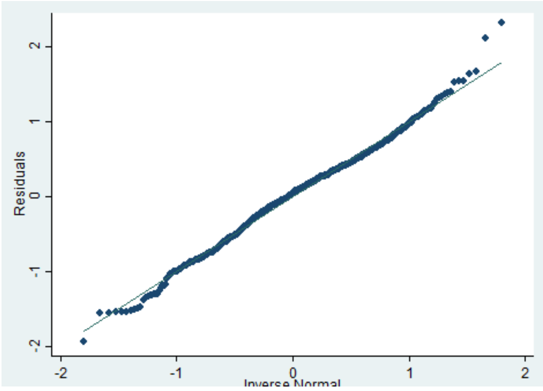


Figure 6.3 Quantiles plot

The figures above show that the deviation of the residuals from the normal distribution is minimal. The Kernel density estimate tells us that there is a slight shift to the right on the residuals, but this is trivial and the error terms can be considered normally distributed. This is supported by the standardized normal probability plot, where it is evident that the residuals follow the straight line closely. The plot of variable quantiles against the quantiles of a normal distribution shows that there are some slight deviations from normal at the tails, as can be seen in the Kernel density estimate above. The *qnorm* is sensitive to non-normality near the tails, and we can expect to see greater deviations from normal here than in the mid-range. In total, the deviations from normal seem to be minor and trivial, and the assumption of normality is considered upheld.

6.1.2 Collinearity

To ensure that the estimations of the coefficients remain stable and that the error terms do not increase, it is necessary to control for collinearity/multicollinearity. A correlations matrix like the one displayed in table 6.1 gives some indications of whether collinearity will be problematic or not, but it is common practise to conduct a VIF-test on the regression to confirm the indications from the correlations matrix. The variance inflation factor (VIF) provides an index that measures how much the variance of an estimated regression coefficient is increased due to collinearity.

Variable	VIF	1/VIF
lnpar30	1.21	0.827276
sqrt_age	1.19	0.842344
DM_individ	1.15	0.870345
lnloan_dis~v	1.12	0.892121
rural	1.04	0.959996
lnTotalass~j	1.03	0.970103
lninflation	1.01	0.986814
Mean VIF	1.11	

Table 6.2 VIF-test

General levels of multicollinearity tolerance are up to 0.10, which corresponds to a VIF of 10 (Hair et al., 2010). Overall, the variables included in this model show satisfying results in the VIF-test, and the underlying assumption for multiple regression is met.

6.1.3 Independence of error terms (autocorrelation)

To control for autocorrelation a paired comparison of the variables related to the same variables in the previous year can be conducted.

	lnoper~s	L. lnoper~s
lnoperexp_~s		
--.	1.0000	
L1.	0.9097	1.0000
lnpar30		
--.	-0.1400	-0.1799
L1.	-0.1155	-0.1598
lnloan_dis~v		
--.	-0.4963	-0.4735
L1.	-0.5193	-0.5036
lnTotalass~j		
--.	-0.4251	-0.3591
L1.	-0.4365	-0.3791
sqrt_age		
--.	-0.1375	-0.1093
L1.	-0.1401	-0.1076
lninflation		
--.	0.1033	0.0838
L1.	0.0575	0.0550
rural		
--.	-0.1459	-0.1232
L1.	-0.1507	-0.1340

Table 6.3 extract from test for autocorrelation

The extract above illustrates the comparison between the dependent variable in two continuous years. The full report can be found in appendix 3. The results show no specific pattern and a low degree if autocorrelation. Accordingly, the model does not breach this assumption.

6.1.4 Homoscedasticity

To test the basic assumption that the variance of the error term is constant, meaning that there is homogeneity in the variance, one can create a scatter plot of residuals compared to the predicted values. A well-fitted model is recognized by the absence of patterns to the residuals when they are plotted against the fitted values (Chen et al., 2003). In cases where the variance of the residuals is non-constant the residual variance is said to be heteroscedastic.

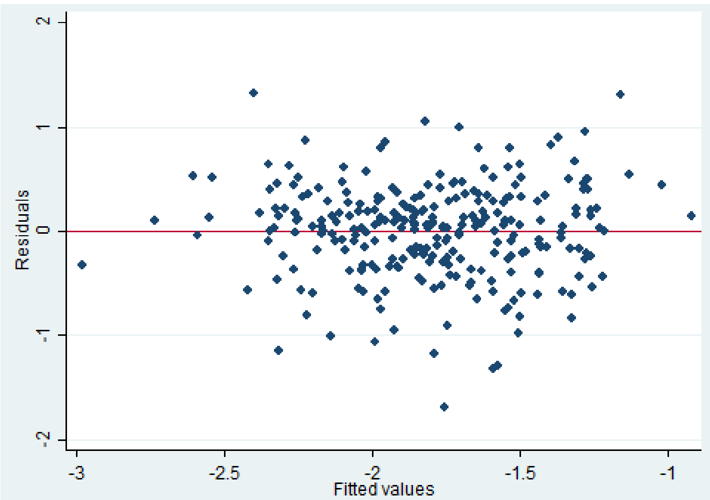


Figure 6.4 Residuals plotted against fitted values

The plot shows some small tendencies towards heteroscedasticity near the red line, which indicates that it may be a problem for the model. Because the tendencies are not severe, it is difficult make any conclusions based on the plot alone. Further tests are necessary to investigate whether the assumption of homoscedasticity is upheld.

To control for heteroscedasticity a White's test will be conducted. The implementation of the test will be based on the results from the VIF-test and thus, it includes the variables with the lowest amount of multicollinearity. The null hypothesis is that there is homoscedasticity in the variance. If the p-values in the White's test are below the chosen 5% significance level, H_0 can be rejected and the alternative hypothesis that there is heteroscedasticity in the variance will remain.

Cameron & Trivedi's decomposition of IM-test

Source	chi2	df	p
Heteroskedasticity	59.32	33	0.0033
Skewness	9.30	7	0.2321
Kurtosis	3.56	1	0.0593
Total	72.18	41	0.0019

Figure 6.5 White's and Cameron & Triverdi's test

The result of the White's test show that prob>chi2 = 0.0033, and the null hypothesis will thus be rejected. This means that based on the White's test there are problems with heteroscedasticity in this model. The test is very sensitive to model assumptions, such as the assumption of normality (Chen et al., 2003). Due to this sensitivity, it is common practice to combine the test with the Breusch-Pagan test and the plot in figure 6.4 to make a judgement on the severity of the heteroscedasticity and to decide if correction is needed.

```
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: fitted values of lnoperexp_assets

         chi2(1)      =     1.20
         Prob > chi2  =   0.2731
```

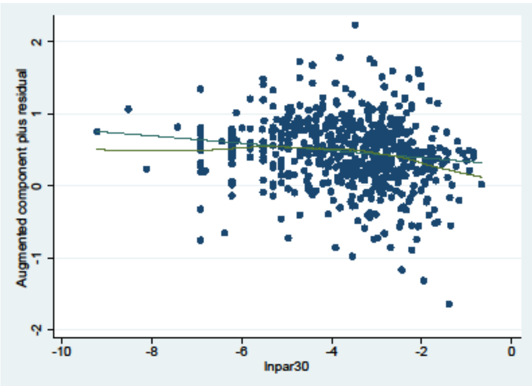
Figure 6.6 Breusch-Pagan test for heteroscedasticity

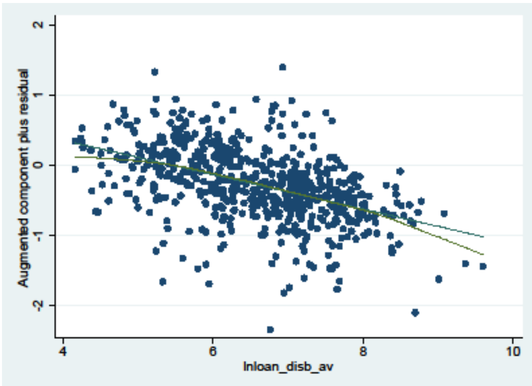
In the Breusch-Pagan test the null hypothesis is that there is a constant variance, meaning that the variance of the residuals is homogeneous. The result of this test gives us a p-value of 0.2731. The null hypothesis is thus not rejected at a 10% significant level. Consequently, the residual plot, White's and Cameron & Triverdi's test and Breusch-Pagan test demonstrate inconsistent results. Because the residual plot does not show severe tendencies of heteroscedastic it could be considered acceptable to go ahead with the analysis without corrections for heteroscedasticity. However, because constant variance of error terms is one of the most important assumptions for multivariate analysis, we will make adjustments to ensure that heteroscedasticity is not a problem in this model.

By using robust standard errors we are able to combat several trivial concerns about failure to meet assumptions, like issues with normality or heteroscedasticity (Chen et al., 2003). When we implement the robust option the point estimates of the coefficients remain the same as in ordinary OLS, but the standard errors consider issues concerning heterogeneity and lack of normality. This is illustrated by changes in the standard errors and t-tests, but will not result in changes in the coefficients. The model will be run with robust standard errors in the analysis starting in chapter 6.1.5.

6.1.5 Linearity

When conducting linear regression, the relationship between the response variable and the predictors should be linear. If this assumption is not withheld, the linear regression will, according to Chen et al. (2003): "try to fit a straight line to data that does not follow a straight line". Conducting a scatter plot between the predictor and the response variable will control for linearity. If nonlinearity is present we can expect to see a plot that does not follow a straight line, e.g. a big wave-shaped curve or a curved band.





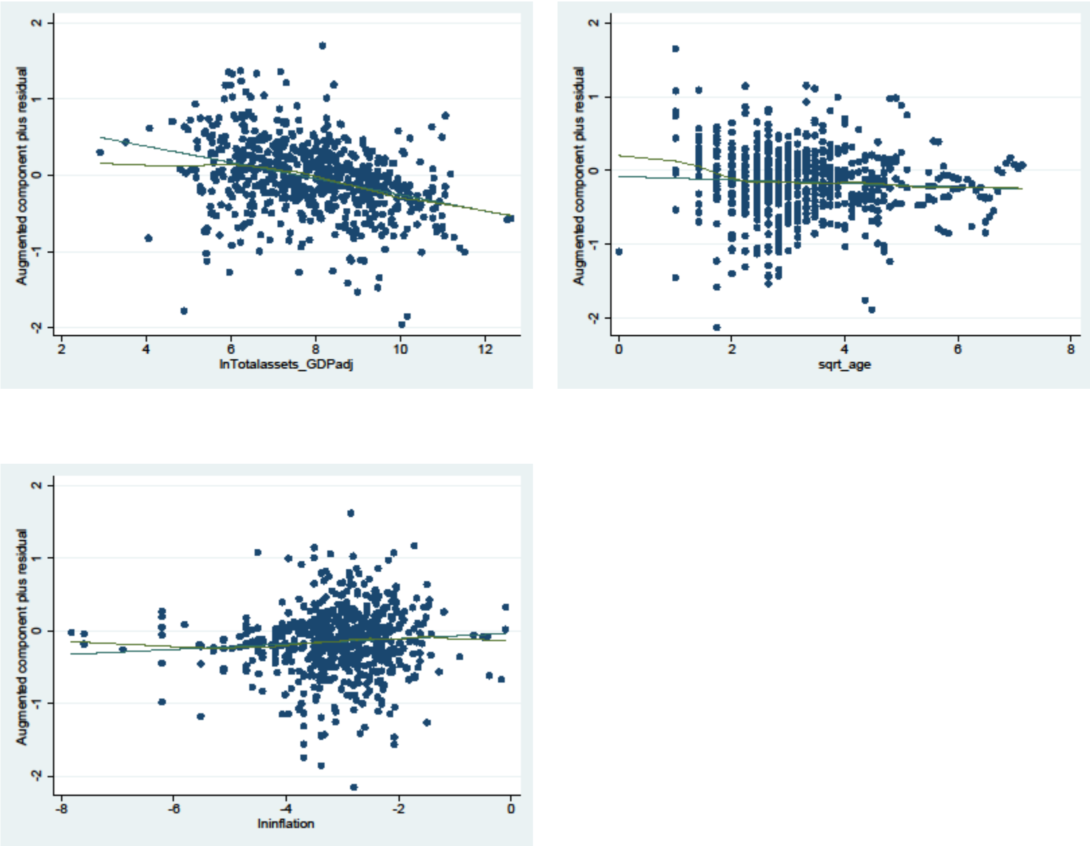


Figure 6.7 Linearity scatter plots

The results of the scatterplots are various, but in this case they are considered to be acceptable in regards to the linearity of the control variables. Adjustments could be made to ensure that the plots more accurately follow the straight line, but such adjustments have proved to aggravate issues in other assumptions, such as normality and multicollinearity. These assumptions are considered to be more important - and the issues associated with them more problematic. Thus, there will be no adjustments to the control variables in the model on the basis of the tests for linearity.

It is difficult to make a statement on the linearity of the independent variable, *lnpar30*, based on its scatterplot. Theoretically it is suspected to have a non-linear effect on operating costs, and thus the relationship will be further investigated in the following section where we test the functional form in the model.

Test of functional form

In order to test the functional form of the model it is necessary to conduct a preliminary regression on which the functional tests will be based.

Linear regression

Number of obs	=	636
F(7, 628)	=	65.54
Prob > F	=	0.0000
R-squared	=	0.3723
Root MSE	=	.46222

lnoperexp_assets	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
lnpar30	-.0484258	.0154136	-3.14	0.002	-.0786943	-.0181573
lnloan_disb_av	-.2416118	.0199881	-12.09	0.000	-.2808634	-.2023603
lnTotalassets_GDPadj	-.1179711	.0133246	-8.85	0.000	-.1441373	-.091805
sqrt_age	-.0240487	.0181413	-1.33	0.185	-.0596736	.0115762
lninflation	.0374997	.0168221	2.23	0.026	.0044652	.0705341
rural	-.3691212	.0576318	-6.40	0.000	-.4822956	-.2559467
DM_individ	-.0916267	.0687826	-1.33	0.183	-.2266985	.0434451
_cons	.9298479	.1848867	5.03	0.000	.5667768	1.292919

Figure 6.8 Preliminary regression with robust standard errors.

The results show a negative and significant effect of default on loans on operating costs per assets. However, it is unclear whether this is modelled correctly, as we might have breached the assumption of linearity. It is suspected that the effect of *lnpar30* is quadric and thus the shape of the bivariate relationship between *lnpar30* and *lnoperexp_assets* should be assessed. A variable that predicts the value based on a locally weighted regression of default on loans on operating costs is created using the lowess command.

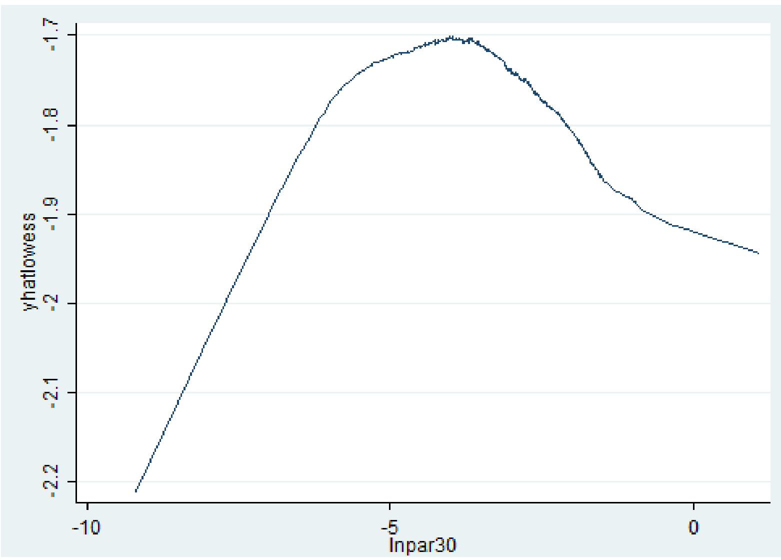


Figure 6.9 Prediction of relationship between lnpar30 and lnoperexp_assets

The graph illustrates a concave curvilinear relationship. This should be modelled this in the regression. When operating with a continuous measure like *lnpar30* the best way of including the relationship in the model is by including a quadratic term (lnpar30*lnpar30) in our model, presented by *c.lnpar30#c.lnpar30* in the table below.

```
Linear regression                               Number of obs     =        693
                                                F(8, 684)         =      60.38
                                                Prob > F          =     0.0000
                                                R-squared         =     0.3590
                                                Root MSE          =      .4661
```

lnoperexp_assets	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
lnpar30	-.24371	.0630051	-3.87	0.000	-.3674167	-.1200033
c.lnpar30#c.lnpar30	-.022212	.0071851	-3.09	0.002	-.0363195	-.0081044
lnloan_disb_av	-.210486	.0199707	-10.54	0.000	-.2496972	-.1712749
lnTotalassets	-.1426609	.0171551	-8.32	0.000	-.1763439	-.108978
sqrt_age	.0008073	.0188551	0.04	0.966	-.0362135	.037828
lninflation	.0440325	.0166293	2.65	0.008	.0113819	.0766832
rural	-.4219639	.0546052	-7.73	0.000	-.5291778	-.31475
DM_individ	-.0997481	.0690443	-1.44	0.149	-.2353123	.0358161
_cons	1.555826	.2701207	5.76	0.000	1.02546	2.086191

Figure 6.10 Regression including quadric term for dependent variable

Based on the above regression the plot of the effect of default on loans on operating costs is extracted:

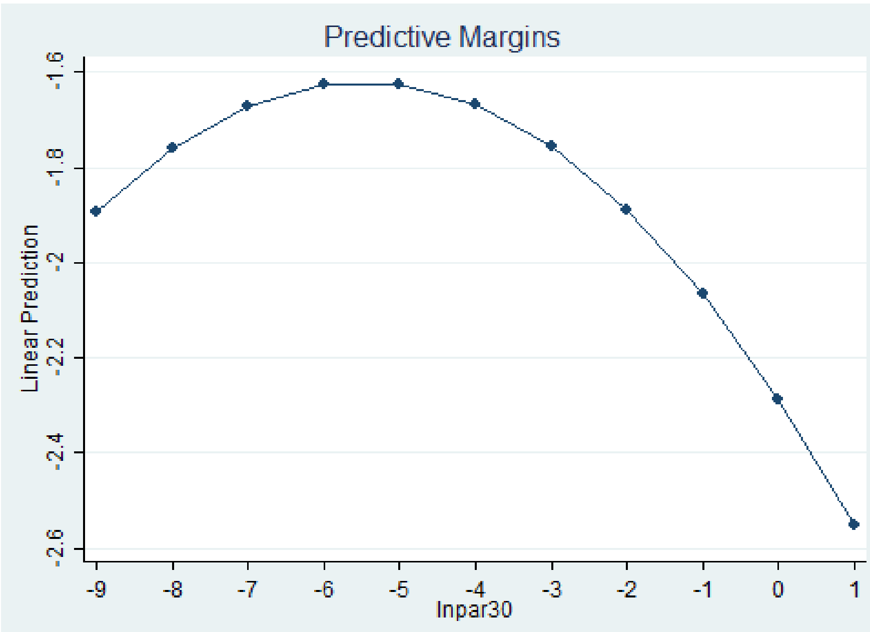


Figure 6.11 Predictive margins

The results tell us that there is a quadric and concave relationship between default on loans and operating costs in microfinance institutions, and the model will thus include the term for the quadric function of loan default.

6.2 Fixed or random effect

To decide whether to go for a model with fixed or random effects in the data analysis, a Hausman test will be conducted. Theory tells us that if there is reason to believe that differences across entities have some influence on the dependent variable, random effects should be used (Torres-Reyna, 2007). It is reasonable to assume that individual characteristics in the microfinance institutions are correlated to operating costs in each institution, so according to theory, the random effects method should be selected for the purposes of this analysis. In contrast to the fixed effects model the random effects model further assumes that the variation across the microfinance institutions are random and uncorrelated with the rate of default on loans. To verify this, a Hausman test will be executed. Firstly, the dataset will be declared panel data:

```
xtset case yr

       panel variable:  case (unbalanced)

        time variable:  yr, 1996 to 2012

                delta:  1 unit
```

Figure 6.12 Panel- and time variable specification

The figure confirm that the dataset is declared to be panel data, where the panel variable is the case, which is the identification variable for the different microfinance institutions (one case = one microfinance institution). The dataset is set up to sort the microfinance institutions after their case number and after the year of which the data was retrieved. Moreover, figure 6.12 display the dataset to be categorized as unbalanced, meaning that not all individual units (cases) are observed in all time periods (years) ($T_i \neq T$ for all i). Panel data theory confirms that this will not be an issue in the analysis.

To conduct the Hausman test, it is necessary to carry out one regression using fixed effects and one regression using random effects. The results of these regressions can be found in appendix 4. After the fixed- and random effects regressions are run, the Hausman test is conducted to test which model is appropriate in this study.

```
                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |     fixed        random       Difference          S.E.
-------------+----------------------------------------------------------------
     lnpar30 |    .0205406    -.1700056        .1905463        .1212639
   c.lnpar30#|
   c.lnpar30 |    .0025399    -.0142405        .0167803        .0100934
 lnloan_dis~v |   -.0730637    -.1918783        .1188146         .087802
 lnTotalass~s |   -.2920941    -.1432453       -.1488488         .090721
     sqrt_age |     .246378    -.0122973        .2586753        .2146714
  lninflation |    .1643654     .0739727        .0903927        .0585327
        rural |   -.0659818    -.4336564        .3676746        .1923289
    DM_individ |    .1719537       .03186        .1400936        .2395585
------------------------------------------------------------------------------
                          b = consistent under Ho and Ha; obtained from xtreg
           B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(8) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =       14.43
                Prob>chi2 =      0.0712
```

Figure 6.13 Hausman test

In the Hausman test the null hypothesis is that the difference in coefficients is not systematic, meaning that random effects should be used. The alternative hypothesis is that the difference in coefficients is systematic and that fixed effects should be used. The test shows *prob>chi2 = 0.0712.* At a chosen significant level of 5% H_0 can thus not be rejected. Consequently, the results of the Hausman test in figure 6.13 support the theory that random-effects is appropriate for this study.

6.2.1 Random effects results

The random effects regression provides the following results:

```
Random-effects GLS regression                   Number of obs      =       693
Group variable: case                            Number of groups   =       237

R-sq:                                           Obs per group:
     within  = 0.1603                                         min =          1
     between = 0.3061                                         avg =        2.9
     overall = 0.3349                                         max =          8

                                                Wald chi2(8)       =    181.09
corr(u_i, X)   = 0 (assumed)                    Prob > chi2        =    0.0000

------------------------------------------------------------------------------------
    lnoperexp_assets |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
---------------------+--------------------------------------------------------------
             lnpar30 |  -.1291724   .0453336    -2.85   0.004    -.2180247   -.0403201
                     |
 c.lnpar30#c.lnpar30 |  -.0104294   .0049484    -2.11   0.035    -.0201281   -.0007308
                     |
     lnloan_disb_av  |  -.1226191   .0236638    -5.18   0.000    -.1689993   -.0762388
       lnTotalassets |  -.1504919   .0197396    -7.62   0.000    -.1891808    -.111803
            sqrt_age |   .0147888   .0266389     0.56   0.579    -.0374224        .067
         lninflation |   .0334025   .0145175     2.30   0.021     .0049488    .0618563
               rural |  -.3453539   .0707813    -4.88   0.000    -.4840828   -.2066251
          DM_individ |   .0312962   .0467832     0.67   0.504    -.0603972    .1229897
               _cons |   1.096922   .2581478     4.25   0.000      .590962    1.602883
---------------------+--------------------------------------------------------------
             sigma_u |    .461242
             sigma_e |   .2157039
                 rho |  .82054326   (fraction of variance due to u_i)
------------------------------------------------------------------------------------
```

Figure 6.14 Random-effects regression

When evaluating the signs of the coefficients we see that most of them are consistent with the assumptions made in chapter 4 about the effects of these variables. Default on loans (*lnpar30)* negatively influences operating costs (*lnoperexp_assets*), suggesting that as the default on loans decreases, the operating cost rate will increase. As default on loans constitute the main explanatory variable in this study, it is important that this is confirmed significant. At a 5% significant level, the p-value of 0.004 states that the effect of the variable on operating costs is in fact significant.

The variable *c.lnpar30##c.lnpar30* represents the squared term of the independent variable, default on loans. This variable also negatively influence the operating costs per assets, and by a 5% significant level it is indeed significant.

The negative sign of the loan size variable, *ln_loan_disb_av*, is consistent with the assumptions made about it; the smaller the loans are, the higher the microfinance institutions' costs. This can be explained by the assumption that the costs of processing a loan will be similar regardless of the loan size, leaving small loans to be relatively more costly. This control variable can be considered significant at a very low significant level and the effect of it on the operating costs rate is noteworthy.

The microfinance institutions' size, measured by the GDP-adjusted total assets, *lnTotalassets_GDPadj*, is shown to be negatively correlated with the operating costs rate of the institutions, suggesting that the microfinance institutions' scale increases, the operating cost rate will decrease. This is consistent with the theory of scale advantages, discussed in chapter 4. At a 5% significant level also this effect on operating costs per assets is significant.

Based on this regression the age of the microfinance institutions, *sqrt_age*, is positively correlated to the operating costs rate of the institution, suggesting that as the microfinance institutions age, the operating costs rate in the will increase. This is not consistent with theory about "learning-by-doing" and the expectation of a negative relationship. One possible explanation is that routines and procedures may not be as tightly followed over time, i.e. that the institutions experience "slack" that outweigh the effect of gained experience. However, at a 5% or 10% significant level the effect of the microfinance institution's age on the operating costs is not significant in the model.

The model shows that inflation will have a positive effect operating costs, which means that the cost will increase with the general inflation in the area where microfinance institution is located. This is consistent with the theoretical expectations, and the effect can be said to be significant on a 5% significant level.

The data output shows that microfinance institutions located in a rural area will negatively influence the operating costs per assets of that institution, suggesting that financing in rural areas is in fact cost-efficient. Consequently, the results of our regression contradict our expectations of a positive relationship. This result may be explained by Wydick (1999)'s findings that compared to urban groups, the rural groups are more willing to apply social pressure to ensure repayment. Further research is however needed in order to confirm or reject this assertion. Other explanations may be that monitoring costs are lower in rural areas, as unconventional methods often are used here (less bureaucracy), or that members of rural communities take pride in being able to handle their finances and business opportunities. Further research is suggested to draw any conclusion about the reasons for the negative relationship between outreach to rural areas and the operating costs rate in microfinance institutions. The variable can be shown to be significant at very low significant levels.

The dummy variable *DM_individ* controls for the lending methodology of the microfinance institutions, and whether or not this will impact the operating costs of the microfinance institution. The relationship is shown to be positive, suggesting that as the level of loans to individuals increases, the operating costs will also increase. Accordingly, our regression supports the arguments that the benefit that comes with group lending leads to reduced costs. At a 10% significant level the effect on the dependent variable is however not significant.

Variable	**Significant at a 10% level**	**Effect on dependent variable**
Independent variables		
lnpar30	Significant	Negative
lnpar30_sqrt	Significant	Negative
Control variables		
Loan size	Significant	Negative
MFI size	Significant	Negative
MFI age	Not significant	Positive
Inflation	Significant	Positive
Market orientation	Significant	Negative
Loan methodology	Not significant	Positive

Table 6.4 Overview of variables, their significance and effect on dependent variable

6.3 Endogeneity

The purpose of this study is to determine whether the level of default on loans the microfinance institutions undertake will have an impact on the operating costs measured by operating costs per assets, and whether the relationship between these to variables exists in that order. As discussed in chapter 4, theory provides reason to believe that the independent variable, *lnpar30*, is endogenous.

The employment of instrumental variables is a widely used strategy when dealing with endogeneity (Bound, Jaeger & Baker, 1995). The instrument variable should not be directly associated with the outcome and should thus have a low correlation with the dependent variable and high correlation with the independent variable, which is suspected to be endogenous (Bound et al., 1995). For this study lagged variables of the default on loans rate will be used as instrument variables to control for endogeneity. It is reasonable to expect that microfinance institutions with high levels of default one year also had somewhat high levels in previous or following years, as changes in the practice concerning collection and repayment are typically implemented over time. At the same time, last years- or the year before that- levels of default are not suspected to be directly correlated to this year's operating costs.

6.3.1. Test for endogeneity

When conducting instrumental variables regressions it is interesting to perform a test of endogeneity to investigate whether or not the explanatory variable in fact is endogenous. This test, as well as the following tests for weak instruments and overidentification can only be run on the basis of a 2SLS instrumental variable regression, i.e. before we adjust for fixed or random effects. Thus, the results for the 2SLS IV-regression is displayed below:

```
Instrumental variables (2SLS) regression          Number of obs   =        438
                                                  Wald chi2(8)    =     296.93
                                                  Prob > chi2     =     0.0000
                                                  R-squared       =     0.3725
                                                  Root MSE        =      .4254
```

lnoperexp_as~s	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
lnpar30	-.358502	.1720553	-2.08	0.037	-.6957243	-.0212798
lnpar30_sqrt	-.0345323	.0193881	-1.78	0.075	-.0725323	.0034676
lnloan_disb_av	-.1929479	.0249526	-7.73	0.000	-.2418541	-.1440418
lnTotalassets	-.1539115	.0202861	-7.59	0.000	-.1936715	-.1141515
sqrt_age	.0024967	.023422	0.11	0.915	-.0434096	.048403
lninflation	.0591161	.0216649	2.73	0.006	.0166537	.1015786
1.rural	-.4112398	.0584862	-7.03	0.000	-.5258707	-.2966089
1.DM_individ	-.0292372	.0828367	-0.35	0.724	-.1915942	.1331198
_cons	1.37097	.423723	3.24	0.001	.5404886	2.201452

```
Instrumented:  lnpar30
Instruments:   lnpar30_sqrt lnloan_disb_av lnTotalassets sqrt_age lninflation
               1.rural 1.DM_individ lag1par30 lag2par30
```

Figure 6.15 Results of 2SLS IV-regression

Based on these results the test for endogeneity can be conducted. If the variable is endogenous it means that there is a reversed causality between the level of default and the operating costs rate in microfinance institutions. The null hypothesis is the independent variable (*lnpar30*) is exogenous. If the P-values for the Durbin (score) statistics and the Wu-Hausman statistics are low, the null hypothesis can be rejected, i.e. the explanatory variables are endogenous. The results of the test for endogeneity are reported in figure 6.16 below:

```
Tests of endogeneity
Ho: variables are exogenous

Durbin (score) chi2(1)          =  .350256  (p = 0.5540)
Wu-Hausman F(1,428)             =  .342533  (p = 0.5587)
```

Figure 6.16 Results of test for endogeneity

Here we see that both the P-values are above both the 5% and 10% significant level, meaning that we cannot reject the hypothesis that *lnpar30* is exogenous. The implications this cause for the study is that we are torn between the theoretical arguments of endogeneity and the results of the test displaying that it will likely not be an issue in our model. Because the theoretical evidence (see Berger & De Young,

1997) tell us that the default on loans rate should be treated as an endogenous variable, we will go a head with the additional test related to endogeneity despite the results of the endogeneity test.

6.3.2 Test for weak instruments

This test is conducted to control whether the instruments chosen are good for the model. When testing for weak instruments we are interested in the correlation between the instruments and the (suspected) endogenous variables. Here we look at the partial R^2, which measures the correlation between the default on loans (*lnpar30*) and the lagged variables of the default on loans when we have eliminated the effect of the exogenous variables (the control variables). This correlation should be high. The other thing we are interested in is our F-statistic. Here the null hypothesis is that the instruments are weak. If the p-value is low it means that the null hypothesis can be rejected and that the instruments are not weak.

First-stage regression summary statistics

Variable	R-sq.	Adjusted R-sq.	Partial R-sq.	F(2,428)	Prob > F
lnpar30	0.9654	0.9646	0.2067	55.7599	0.0000

Figure 6.17 Results of test for weak instruments

The results show a partial R^2 of 0.2067, which is not high, but it is very low either. The p-value (Prob>F) is shown to be 0.0000, which means that the null hypothesis can be rejected: the instruments used are not weak and can be considered good for the model.

6.3.3 Test for overidentification

This test is carried out to control the number of instruments compared to the number of endogenous variables to see if the model is correctly specified. What we want to see is overidentification, which means that the number of instruments is greater than the number of endogenous variables. Justified identification, where the number of instruments equals the number of endogenous variables, is also acceptable. What we want to avoid is underidentification, where there are less instrumental variables than endogenous variables.

In the test, the null hypothesis is that the instrument set is valid and the model is correctly specified. We are again interested in the p-value, which should be high to confirm that the model is correctly specified. If the levels are lower than 0.05, the null hypothesis has to be rejected and the model cannot be said to be correctly specified.

```
Tests of overidentifying restrictions:

Sargan (score) chi2(1) =  .747549  (p = 0.3873)
Basmann chi2(1)        =  .731731  (p = 0.3923)
```

Figure 6.18 Results of test for overidentification

The results of our test of overidentification show a high p-value of 0.3873, indicating that the model is correctly specified.

6.3.4 Results of IV-regression using random effects

The model displayed in figure 6.19 below is carried out using random effects, as supported by the Hausman test conducted in chapter 6.2. Because of the results of the endogeneity test in chapter 6.3.1 we chose to display the results ignoring the instrumental variables and treating all variables as exogenous. The instrumental variables regression (IV-regression) show that the default rate on loans still significantly influences the operating costs rate, both in terms of the default on loans rate and the squared function of it. By utilizing instrument variables regression, we are able to hold the explanatory variable (*lnpar30* and *lnpar30_sqrt*) virtually exogenous and not endogenous. Any problems concerning endogeneity can thus be considered to be relieved.

```
G2SLS random-effects IV regression              Number of obs      =       438
Group variable: case                            Number of groups   =       214

R-sq:                                           Obs per group:
     within  = 0.1955                                         min =          1
     between = 0.3413                                         avg =        2.0
     overall = 0.3547                                         max =          6

                                                Wald chi2(8)       =    158.29
corr(u_i, X)       = 0 (assumed)                Prob > chi2        =    0.0000
```

lnoperexp_as~s	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lnpar30	-.1722479	.0601822	-2.86	0.004	-.2902029	-.0542929
lnpar30_sqrt	-.0136516	.0065533	-2.08	0.037	-.0264959	-.0008072
lnloan_disb_av	-.11603	.0256591	-4.52	0.000	-.166321	-.0657391
lnTotalassets	-.1739183	.0225579	-7.71	0.000	-.218131	-.1297055
sqrt_age	.0049442	.0283714	0.17	0.862	-.0506628	.0605512
lninflation	.0553353	.0193141	2.87	0.004	.0174804	.0931902
1.rural	-.2971485	.0725155	-4.10	0.000	-.4392763	-.1550207
1.DM_individ	.0208449	.0570578	0.37	0.715	-.0909863	.1326761
_cons	1.418381	.3161232	4.49	0.000	.7987906	2.037971
sigma_u	.38871221					
sigma_e	.19350556					
rho	.80139977	(fraction of variance due to u_i)				

```
Instrumented:   lnpar30
Instruments:    lnpar30_sqrt lnloan_disb_av lnTotalassets sqrt_age lninflation
                1.rural 1.DM_individ lnpar30
```

Figure 6.19 Results from G2SLS random-effects IV regression

The concave bivariate relationship between the independent and dependent variables (as shown in figure 6.11) tell us that operating costs increase with default on loans initially, but after a point the operating costs will decrease with increasing defaults. This suggests an initial positive effect and a subsequent negative effect of the independent variable. An explanation for this is that when default rates are increasing initially, operating costs will also increase, as more resources are committed to reduce the defaults. Over time the microfinance institutions might become more cost-efficient even though default rates continue to increase. Another alternative is that fixed investments were incurred and that they over time have started providing benefits for the institution. However, the concave curve graphed in the predictive margins is not supported by the multivariate regression results in figure 6.19 because both the linear and squared term of the independent variable has a negative coefficient. In order to conclude that the relationship between default on loans and operating costs is non-linear we wish to see opposite signs of the linear and squared term of the default on loans, and that they both are significant. The concave graph displayed in the predictive margins merely illustrates the bivariate relationship between the two variables. A bivariate analysis is one of the simplest forms of quantitative analysis and is used to determine the empirical relationship between two variables.

Once we include a square term in the regression using with robust standard errors and the results give the same coefficient signs and statistical insignificance, the squared term can be disregarded. The results of the instrumental variable regression show that there is a negative relationship between the default on loans rate (*lnpar30)* and the operating costs rate, suggesting that as risk increases, operating costs will decrease. This finding support Berger and De Youngs (1997) assertion that banks incur additional costs from trying to maintain low levels of default, as it forces the bank to spend extra resources on monitoring and underwriting to influence loan quality. The results suggest that microfinance institutions might gain from looking into making some adjustments to lower the monitoring and supervisory costs that are incurred from maintaining low rates of loan default. Further studies are however recommended to get a clearer picture of the functional form of the relationship, as the multivariate regression in this analysis does not support the findings in the bivariate analysis.

6.4 Reliability and validity

The reliability and validity of a study is concerned the quality of the work. The reliability refers to whether the data can be considered trustworthy/reliable, while the validity is concerned with its relevance to the associated theory and the research objective (Ringdal, 2013).

6.4.1 Reliability

The reliability of the data will be improved if it is possible to achieve similar results in repeated measurements using the same measurement concepts (Ringdal, 2013). This can be tested for by conducting additional regressions using the same control variables, but altering the variables used to measure the concepts of the dependent and independent variable.

Alternative measures

For this study we will use the operating expenses related to portfolio as an alternative measure for the operating cost rate and conduct the same regressions and tests as with the operating expenses related to assets. The alternate variable will be replacing the initial variable in such a way that we get results for the relationship between Par30 and operating expenses related to assets as well as Par30 and operating expenses related to portfolio. As an alternative to the independent variable this study will use the combined credit risk. This variable is a combination of the write off and par30 in the microfinance institutions, and is thus a suitable measure for the risk of default. Regressions will be run using both operating expenses related to assets and operating expenses related to portfolio as the dependent variable and the combined credit risk as the independent variable. Consequently, we attain results from four different combinations of variables, as illustrated in the matrix below. If the alternative analyses show similar results to those presented in the previous chapters, the reliability of the study and the results are strengthened. The control variables will remain the same in all regressions.

	lnpar30	lncomb_credrisk
lnoperexp_assets	Quadric bivariate relationship Negatively correlated Significant at a 5% significant level	Quadric bivariate relationship Negatively correlated Insignificant at a 5% significant level
lnoperexp_portf	Quadric bivariate relationship Negatively correlated Insignificant at 5% and 10% significant level	Quadric bivariate relationship Negatively correlated Insignificant at 5% and 10% significant level

Table 6.5 Results with alternative measures

The table show that all models find a negative, quadric bivariate relationship between the default on loans and the operating costs rate. The full regressions using alternative measures can be found in appendix 3. We do however note that the relationship between the independent and dependent variable is shown to be insignificant when using alternative measures. This strengthens the need for further studies on the relationship between the default on loans rate and operating costs rate in microfinance institutions. On the other hand, the results using alternative measure are consistent and have the same signs on the coefficients in all regressions. Additionally, they all have the same shape of the bivariate relationship between the dependent and independent variable as we see the initial analysis using *par30* and operating expenses related to assets.

6.4.2 Validity

The validity of the study cover the entire experimental concept and is concerned with whether or not it measures what we intended to measure and to what degree you can draw conclusions about the research objective (Braut, 2009). Additionally, it ascertains whether the requirements of scientific research method is met. It is common to talk about both internal and external validity, which respectively concerns the degree to which the experimental design is structured correctly and the variables measure the concept it was meant to (Ringdal, 2013), as well as the process of examining the results and possible causal relationships. In terms of the internal validity it is essential to point out that the purpose of this study was to investigate the relationship between the default on loans rate and the operating costs rate in the microfinance institutions, and that it does not aim to explain, nor explore, the drivers behind the operating costs rate. The control variables in the study are thus included to control that the calculated effect of risk on the operating costs rate in fact stems from that risk variable and not other factors. The regression models portrait the effects of the control variables as well as the independent variable, and the calculated significance of these variables tells us that us that we were correct in including them in the model, as theory suggested. We do however note that the microfinance age and lending methodology did not prove significant, and could thus be left out of the model. On the other hand, there is plenty of theory suggesting that these variables do affect the operating costs of the microfinance institutions. For this reason the variables remained in the model, but we suggest further research as to whether or not they actually affect the dependent variable.

7. Conclusion

Drawing from contract and agency theory, theory on moral hazard, credit risk and Berger & De Youngs (1997) study on cost efficiency, this thesis have argued that there is reason to believe that default rates in microfinance institutions can both positively and negatively affect the operating costs of the institutions. This is argued by discussing how; (i) the effects of utility-maximizing behaviour and that these can be extra prominent in the microfinance industry because the clients often are poor and desperate, (ii) social monitoring procedures of the microfinance industry can actually lower the need for monitoring and controlling measures by the institution itself, (iii) the microfinance institutions incur increased monitoring due to clients lacking credit history and collateral, (iv) moral hazard and information asymmetry leads to evasion of the first-best solution, (v) the microfinance industry has challenges with monitoring due to problems separating good from bad risk, (vi) microfinance institutions can experience difficulties implementing and sustaining credit risk management, (vii) low levels of default also lowers operating costs, as defaults have been shown to leave banks with fewer resources available to for lending to other customers, (iix) banks incur additional costs from loans that do not perform or default because they force the microfinance institution to spend extra resources on monitoring and underwriting to influence loan quality, and (ix) bad management in terms of monitoring and underwriting (i.e. bad credit risk management) will lead to increased operating costs almost immediately, whereas loan defaults typically occur at a later point in time. Existing theory and research makes good arguments for both a positive influence of the default rate on the operating cost rate, as well as a negative influence. These arguments set the basis for investigating the actual relationship between the two variables and whether or not the default on loans rate can actually affect the operating costs both positively and negatively (i.e. a non-linear relationship). The results of this study show that there is a bivariate quadric relationship between default on loans and operating costs, but this is not supported in the multivariate regression as both *lnpar30* and *lnpar30_sqrt* have a negative coefficient. The consequence of this is that we cannot conclude that the relationship between default on loans and operating costs is non-linear, and we do not succeed in illustrating a general functional form for the relationship. Further studies are suggested on this topic as the bivariate results suggests that there might in fact be a non-linear relationship. Also, the study faces some limitations in that the model does not control for the measures microfinance institutions take once they notice that the default on loans is higher than wanted. As argued in previous chapters, many banks incur additional costs from loans that do not perform or default because they feel forced to spend extra resources on monitoring and underwriting to influence loan quality. However, the interesting question regarding the relationship between default on loans and operating costs is how the operating costs would look if the microfinance institution did not try to compensate for the higher default levels by applying more resources in collection and monitoring practice, but in stead settled for a higher level of default. Would they be better off by allowing for higher default levels and consequently, lower costs on monitoring, control and collection practice?

7.1 Suggestions for future studies

Studies controlling for the microfinance response to higher default rates is suggested to improve the quality of the relationship between the default on loans rate and operating costs rate in microfinance institutions. It would be interesting to see how these two variables influence each other when the institutions willingly allow for higher default levels than what has become common in the industry.

This study uses alternative measures of both the dependent and independent variable to check the reliability of the results. Though the results are similar with different variables, they are not significant. I recommend that additional studies using different measurements for default on loans and operating costs are conducted in order to gain a better understanding and more reliable conclusions on the relationship between these two variables.

Our regression analysis shows that there is a negative effect of a rural market orientation and the operating costs in microfinance institutions. This is not aligned with the theoretical expectations. Because there is limited studies conducted on this subject, it would be interesting to investigate the effect of the market orientation on operating costs further.

Generally, any study that explains or explores one or more drivers behind the operating costs rate would be interesting and beneficial in the microfinance sector. As argued in the initial chapters, the industry faces a need for lower operating costs in order to lower the interest rates offered to customers.

8. References

Acock, A. C. (2012). *A gentle introduction to Stata, Rev. 3rd Edition*. Texas: Stata Press.

Adams, D. W., & Fitchett, D. A. (1992). *Informal Finance in Low-Income Countries.* Boulder, CO: Westview Press.

Afriyie, H. O., & Akotey, J. O. (2013). Credit risk management and profitability of rural banks in the Brong Ahafo region of Ghana. *European Journal of Business and Management, 5*(24), 24-34.

Agu, C. C. (1998). Loan Management in Agriculture. In Ijere, M., & Okorie, A. (Eds.), *Readings in Agricultural Finance*. Lagos: Longman Nigeria Plc.

Aguilera, R. V., & Jackson, G. (2003). The cross-national diversity of corporate governance: dimensions and determinants. *Academy of Management Review, 28*(3), 447-465.

Albuntas, Y., Carbo, S., Gardener, E. P. M., & Molyneux, P. (2007). Examining the Relationships between Capital, Risk and Efficiency in European Banking. *European Financial Management, 13*(1). 49-70.

Allen, L., & Rai, A. (1996). Operational efficiency in banking: An international comparison. *Journal of Banking & Finance, 20*(4), 655-672.

Armendariz de Aghion, B. & Morduch, J. (2005). The Economics of Microfinance. Cambridge, MA: MIT Press.

Armendariz de Aghion, B., & Morduch, J. (2010). *The Economics of Microfinance,* 2nd edn. Cambridge, MA: MIT Press.

Banerjee, A., & Duflo, E. (2007). The Economic Lives of the Poor. *Journal of Economic Perspectives, 21*(1), 141-167.

The Basel Committee on Banking Supervision. (2000). *Principles for the Management of Credit Risk*. Retrived from: http://www.bis.org/publ/bcbs75.htm

Bateman, M. (2010). *Why Doesn't Microfinance Work?: The Destructive Rise of Local Neoliberalism.* London: Zed Books.

Berger, A. N., & De Young, R. (1997). Problem loans and cost efficiency in commercial banking. *Journal of Banking & Finance, 21*(6), 849-870.

Bhatt, N., & Tang, S. Y. (2001). Delivering microfinance in developing countries: Controversies and policy perspectives. *Policy Studies Journal, 29*(2), 319-333.

Biswas, S. (2010, 16.12.). India's micro-finance suicide epidemic. *BBC News.* Retrived from: http://www.bbc.com/news/world-south-asia-11997571

Bolton, P., & Dewatripont, M. (2005). *Contract Theory*. Cambridge, MA: MIT Press.

Bound, J., Jaeger, D., & Baker, R., (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association, 90*(430), 443-450.

Braut, G. S. (2009). Validitet. *Store Norske Leksikon*. Retrived 16.06.2016, from: https://sml.snl.no/validitet

Bromiley, P. (2005). *The Behavioural Foundations of Strategic Management*. Malden, MA: Blackwell.

Bruce, A., Buck, T., & Main, B. G. M. (2005). Top executive remuneration: a view from Europe. *Journal of Management Studies 42*(7), 1493–1506.

Business Insider. (2012, 24.2.). Hundreds Of Suicides In India Linked To Microfinance Organizations.

Bøhren, Ø., & Strøm, R. Ø. (2007). Aligned, Informed, and Decisive: Characteristics of Value-Creating Boards. *Ljubljana Meetings Paper*. SSRN: http://ssrn.com/abstract=966407

Cameron, A. C., Trivedi, P. K. (2010). *Microeconomics using Stata, Revised Edition*. Texas: Stata Press.

Caudill, S. B., Gropper, D. M., & Hartarska, V. (2009). Which Microfinance Institutions Are Becoming More Cost Effective with Time? Evidence from a Mixture Model. *Journal of Money, Credit and Banking, 41*, 651–672.

Chakrabarty, S., & Bass, A. E. (2013). Encouraging entrepreneurship: Microfinance, knowledge support, and the costs of operating in institutional voids. *Thunderbird International Business Review, 55*(5), 545-562.

Chen, X., Ender, P. B., Mitchell, M., & Wells, C. (2003). Regression with Stata. Retrieved from: www.ats.ucla.edu/stat/stata/webbooks/reg/chapter2/statareg2.htm.

Christen, R. P., Lyman, T. R., & Rosenberg, R. (2002). *Microfinance Consensus*

Guidelines. Guiding Principles on Regulation and Supervision of Microfinance. Washington, DC: CGAP.

Christen, R. P., Rosenberg, R., & Jayadeva, V. (2004). *Financial institutions with a double-bottom line: implications for the future of microfinance.* Washington, DC: CGAP.

Coyle, B. (2000). Framework for credit risk management. London: Fitzroy Dearborn publishers.

Crémer, J., Khalil, F., & Rochet, J-C. (1998). Contracts and Productive Information Gathering. *Games and Economic Behavior, 25*(2), 174-193.

Cull R., Demigüc-Kunt, A., & Morduch, J. (2007). Financial performance and outreach: A global analysis of leading microbanks. *Economic Journal, 117*(517), 107-133.

Dahlum, S. (2014). Kvantitativ analyse. *Store Norske Leksikon.* Retrieved from: https://snl.no/kvantitativ_analyse

Dehejia, R., Montgomery, H., & Morduch, J. (2012). *Do interest rates matter? Credit demand in the Dhaka slums. Journal of Development Economics, 97*(2), 437–449.

Dembe, A. E., & Boden, L. I. (2000). Moral Hazard: A question of morality?. *New Solutions, 10*(3), 257-279.

Dewatripont, M., & Maskin, E. (1995). Contractual Contingencies and Renegotiation. *RAND Journal of Economics, 26*(4), 704-719.

Di Bella, G. (2011). The Impact of the Global Financial Crisis on Microfinance and Policy Implications. International Monetary Fund, IMF Working Papers: 11/175.

Dichter, T. W., Harper, M. (Eds.). (2007). *What's wrong with microfinance?.* Essex: Practical Action Publishing.

Dowd, K., Bartlett, D.L., Chaplin, M., Kelliher, P., & O'Brien, C. (2008), Risk management in the UK insurance industry: the changing state of practice. *International Journal of Financial Services Management*, *3*(1), 5–23.

Eisenhardt, K. M. (1989). Agency Theory: An Assessment and Review. *The Academy of Management Review 14(*1), 57-74.

Fama, E. F., & Jensen, M. C. (1983). Separation of ownership and control. *The*

Journal of Law and Economics, 26(2), 301-325.

Fernando, N. A. (2006). *Understanding and dealing with high interest rates on microcredit: A note to policy makers in the Asia and Pacific region.* Manila: Asian Development Bank.

Field, E., & Pande, R. (2008). Repayment Frequency and Default in Microfinance: Evidence from India. *Journal of the European Economic Association, 6*(2-3), 501-509.

Fligstein, N., & Freeland, R. (1995). Theoretical and comparative perspectives on corporate organization. *Annual Review of Sociology, 21*.21-43.

Garmaise, M. J. & Natividad, G. (2010). Information, the cost of credit, and operational efficiency: An empirical study of microfinance. *Review of Financial Studies 23*(6), 2560– 2590.

Ghate, P. (2007). Learning from the Andhra Pradesh crisis. In T. Dichter & M. Harper (Eds.), *What's wrong with microfinance?*. Rugby, UK: Practical Action Publishing.

Gonzalez, A. (2007). Efficiency Drivers of Microfinance Institutions (MFIs): The Case of Operating Costs. *MicroBanking Bulletin, 15,* 37-42.

Gonzalez, A. (2010). Analyzing Microcredit Interest Rates: A Review of the Methodology Proposed by Mohammed Yunus. *MIX Data Brief, (4)*. Retrived from: https://www.themix.org/publications/mix-microfinance-world/2010/03/analyzing-microcredit-interest-rates-review-methodology-

Green, C. J., Kirkpatrick, C. H., & Murinde, V. (2005). How does finance contribute to the development process and poverty reduction?. In C. J. Green, C. H. Kirkpatrick, & V. Murinde (Eds.), *Finance and Development: Surveys of Theory, Evidence and Policy.* Cheltenham, U.K.: Edward Elgar.

Greene, W. H. (2012). *Econometric analysis, 7th Edition.* Upper Saddle River, N. J.: Pearson Prentice Hall

Hair, J. F., Black, W.C., Babin, B. J.,& Anderson, R. E. (2010). *Multivariate Data Analysis: A global perspective.* Upper Saddle River, N. J.: Pearson Education

Hair, J. F., Money, A. H., Samouel, P., & Page, M. (2007). *Research Methods for Business.* Chichester, U.K.: John Wiley and Sons

Harris, J. & Bromiley, P. (2007). Incentives to cheat: the influence of executive compensation and firm performance on financial misrepresentation.

Organization Science, 18(3), 350-367.

Hart, O., Holmström, B. (1987). The theory of contracts. In Bewley, T (Ed.). *Advances in Economics and Econometrics* (71-155). Cambridge: Cambridge University Press.

Hartarska, V., Shen, X., & Mersland, R. (2013). Scale economies and input price elasticities in microfinance institutions. *Journal of Banking & Finance*, 37. 118-131.

Hausman, J. A. (1978). Specification Tests in Econometrics. *Econometrica 46*(6), 1252-1271.

Helms, B & Consultative Group to Assist the Poorest (2006). *Access for all: building inclusive financial systems*. World Bank; London: CGAP

Hulme, D., & Mosely, P. (1996b). *Finance against poverty.* Vol. 2. London: Routledge.

Hulme, D. (2000). Is microdebt good for poor people? A note on the darkside of microfinance. *Small Enterprise Development, 11*(1), 26–28.

Investing Answers. (2016). *Credit risk.* Retrived from: http://www.investinganswers.com/financial-dictionary/debt-bankruptcy/credit-risk-2933.

Jain, S., & Mansuri, G, (2003). A little at a time: The use of regularly scheduled repayments in microfinance programs. *Journal of Development Economics, 72*(1), 253-279.

Jansson, T. 2003. Financing Microfinance. Inter-American Development Bank Sustainable Development Department Technical Paper Series, MSM-118, Washington DC.

Jenkins, M. W. (2009). *Essays on consumer credit markets* (doctoral thesis, Stanford University). Retrieved from: https://purl.stanford.edu/yb246bk3096

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics, 3*(4), 305-360.

Joyner, E. (Ed.). (2013). *Business Research Methods, 9th Edition.* Mason, OH: South-western, Cengage learning

Kang, S. H., & Sivaramakrishnan, K. (1995). Issues in testing earnings management and an instrumental variable approach. *Journal of accounting Research*, *33*(2), 353-367.

Karlan, D., & Zinman, J. (2009). Observing Unobservables: Identifying Information Asymmetries With a Consumer Credit Field Experiment. *Econometrica*, *77*(6), 1993-2008.

Köszegi, B. (2014). Behavioral Contract Theory. *Journal of Economic Literature, 52*(4), 1075-1118.

Krugman, P. (2009). *The return of depression economics and the crisis of 2008.* New York: W. W. Norton & Company, Inc.

Laffond, J-J., Martimort, D. (2002). *The Theory of incentives: The Principal-Agent model.* Princeton, N.J.: Princeton University Press.

Laffont, Jean-Jacques; Martimort, David (2002). *The theory of incentives: The principal-agent model*. Princeton University Press.

Levine, R. (2005). Finance and growth: Theory and Evidence. In P. Aghion, & S. N. Durlauf (Eds.), *Handbook of Economic Growth* (865-934). Amsterdam: Elsevier.

Lindgren, H. (1987). Bank, Investment Company, Banking Firm. *Stockholms Enskilda 37 Bank.* Stockholm: Institute for Research in Economic History at the Stockholm School of Economics.

Littlefield, E, Morduch, J., & Hashemi, S. (2003). *Is microfinance an effective strategy to reach the millennium development goals?. Focus Note No. 24.* Washington, DC: CGAP.

MacNeil, I. R. (2000). Rationl contract theory: challenges and queries. *Northwestern University Law review, 94, 877-907.*

Mas-Colell, A., Whinston, M. D, & Green, J.R. (1995). *Microeconomic Theory.* Oxford: Oxford University Press.

Maskin E, & Tirole J. (1992). The Principal-Agent relationship with an informed principal: II: Common values. *Econometrica*, *60*(1), 1-42.

McKinsey&Company. (2016). *Credit risk*. Retrived from: http://www.mckinsey.com/business-functions/risk/how-we-help-clients/credit-risk.

Mersland, R. (2009). The Cost of Ownership in Microfinance Organizations. *World Development*, *37*(2), 469-478.

Mersland, R. & Strøm, R. Ø. (2008). Performance and trade-offs in Microfinance Organisations—Does ownership matter?. *Journal of International Development, 20,* 598–612.

Mersland, R.,& Strøm, R. Ø. (2010). Microfinance Mission Drift. *World Development, 28,* 617-629.

Mersland, R., & Strøm, R. Ø. (2012a). The past and future of innovations in microfinance. In D. Cumming (Ed.), *The Oxford Handbook of Entrepreneurial Finance* (859-891). Oxford: Oxford University Press.

Mersland, R., & Strøm, R. Ø. (2012b). What drives the microfinance lending rate?. *Midwest Finance Association 2013 Annual Meeting Paper.* SSRN: http://ssrn.com/abstract=2144618 or http://dx.doi.org/10.2139/ssrn.2144618

Mersland, R., & Strøm, R. Ø. (2013). Microfinance: Costs, Lending Rates and Profitability. In G. Caprio (Ed.). *Handbook of key global financial markets, institutions and infrastructure.* 489-500. London: Elsevier

Mersland, R. & Urgeghe, L. (2013), International Debt Financing and Performance of Microfinance Institutions. Strat. Change, 22: 17–29. doi:10.1002/jsc.1919

Milgrom, P., & Roberts, J. (1992). *Economics, organization & management*. New Jersey: Prentice Hall.

Montgomery, R. (1996). Disciplining or protecting the poor? Avoiding the social costs of peer pressure in micro-credit schemes. *Journal of International Development, 8(2),* 289–305.

Morduch, J., (1999). The Microfinance Promise. *Journal of Economic Literature, 37*(4), 1569-1614.

Morduch, J. (2000). The microfinance schism. *World Development, 28(*4), 617–629.

Natilson, N, & Bruett, T. A. (2001). Financial Performance Monitoring: A Guide for Board Members of Microfinance Institutions. World Bank; Washington DC: CGAP.

Nobel Media. (2016). *The Nobel Peace Price 2006.* Retrived 20.9.2016 from: https://www.nobelprize.org/nobel_prizes/peace/laureates/2006/

Opportunity International. (2016). *About us.* Retrived from: http://opportunity.org/about-us/

Padmanabhan, K.P. (1988). Rural Credit: Lessons for Rural Bankers and Policy Makers. London: Intermediate Technology Publication Ltd.

Pritchard, J. (2016, 05.08). Moral Hazard - What It Is and How It Works. *The Balance.* Retrieved from https://www.thebalance.com/moral-hazard-what-it-is-and-how-it-works-315515

Rhyne, E. (1998). The Yin and Yang of microfinance: Reaching the poor and sustainability. *MicroBanking Bulletin*, *2*, 6-9.

Rhyne, E. (2011, 25.5.). On Microfinance: Who's to Blame for the Crisis in Andhra Pradesh?. *Huffington Post.* Retrived from: http://www.huffingtonpost.com/elisabeth-rhyne/on-microfinance-whos-to-b_b_777911.html

Rogerson, W. P. (1985). Repeated moral hazard. *Econometrica*, *53*(1), 69-76.

Rosenberg, R. (2007). *CGAP Reflections on the Compartamos Initial Public Offering: A case study on microfinance interest rates and profits.* CGAP Focus Note; no. 42. Washington DC: World Bank.

Rosenberg, R., Gaul, S., Ford, W., & Tomilova, O. (2013). *Microcredit interest rates and their determinants.* CGAP Access to Finance Forum; no. 7. Washington DC: World Bank..

Rosenberg, R., Gonzalez, A., & Narain, S. (2009). *The new moneylenders: Are the poor being exploited by high microcredit interest rates?* CGAP Occasional Paper; no. 15. Washington, DC: World Bank.

Santomero, A. M., & Babbel, D. F. (1997). Financial Risk Management by Insurers: An Analysis of the Process. *The Journal of Risk and Insurance*, *64*(2), *231-270*

SAS. (2016). *Credit risk management: what it is and why it matters*. Retrived from: http://www.sas.com/en_au/insights/risk-fraud/credit-risk-management.html.

Sebstad, J., & Cohen, M. (2000). Microfinance, Risk Management and Poverty. World Bank; Washington DC: working paper.

Simkovic, M. (2016). What can we learn from credit markets? Proceedings of the 93rd Annual Meeting of the American Law Institute, Washington, D.C. *Seton Hall Public Law Research Paper*. Retrived from: http://ssrn.com/abstract=2782844.

Schmitz, P. W. (2005). Allocating control in agency problems with limited liability and sequential hidden actions. *RAND Journal of Economics, 36*(2), 318-336.

Studenmund, A. H. (2011). *Using econometrics: a practical guide, 6th Edition.* Essex, U.K.: Pearson Education Limited.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using Multivariate Statistics,* Boston, MA: Pearson Education Limited.

Tamhane, A. C., & Dunlop, D. D. (2000). *Statistics and Data Analysis: from elementary to intermediate,* Upper Saddle River, N.J.: Pearson Prentice Hall.

Tirole, J. (1988). *The Theory of Industrial Organization*. Cambridge, MA: MIT Press.

Tirole, J. (2006). *The theory of corporate finance.* Princeton, N. J.: Princeton University Press.

Torres-Reyna, O. (2007). Panel Data Analysis, Fixed and Random Effects using Stata. Retrieved 08.03.2016 from http://www.princeton.edu/~otorres/panel101.pdf

Tosi, H. L., Werner, S., Katz, J. P., & Gomez-Mejia, L. R. (2000). How much does performance matter? A meta-analysis of CEO pay studies. *Journal of Management, 26*(2), 301-339.

United Nations. (2006). *Building inclusive financial sectors for development.* New York: United Nations.

United Nations. (2016). *International year of microcredit 2005.* Retrived 20.9.2016 from: http://www.un.org/en/events/pastevents/microcredit_2005/

Varian, H. R. (1990). Monitoring agents with other agents. *Journal of Institutional and Theoretical Economics, 146*(3), 153-174.

Verbeek, M. (2012). *A guide to modern econometrics, 4th Edition.* Chichester, U.

John Wiley & Sons Limited.

Wiseman, R. M., Cuevas-Rodríguez, G., & Gomez-Mejia, L. R. (2012). Towards a Social Theory of Agency. *Journal of Management Studies 49(*1), 202-222.

Westphal, J. D., & Khanna, P. (2003). Keeping directors in line: social distancing as a control mechanism in the corporate elite. *Administrative Science Quaterly, 48.* 361-398.

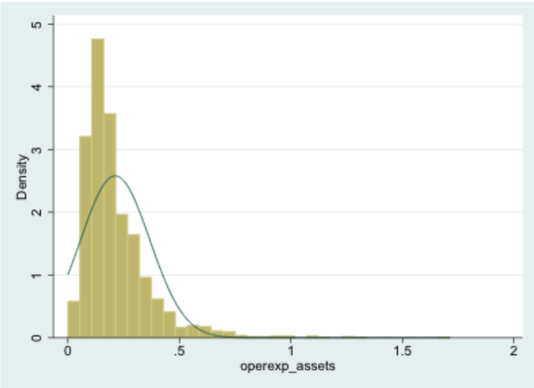
Wydick, B. (1999). Can social cohesion be harnessed to repair market failures? Evidence from group lending in Guatemala. *The Economic Journal, 109*(457), 463-475.

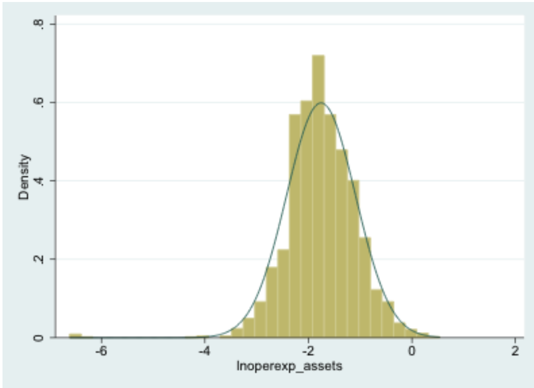
Yunus, M.,& Jolis, A. (1999). *Banker to the poor: micro-lending and the battle against world poverty*. New York: PublicAffairs.

Appendix

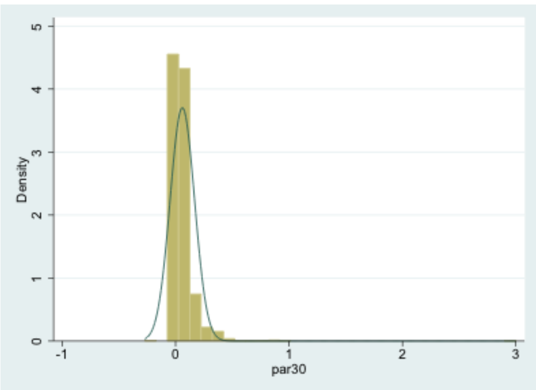
Appendix 1: Distribution tests

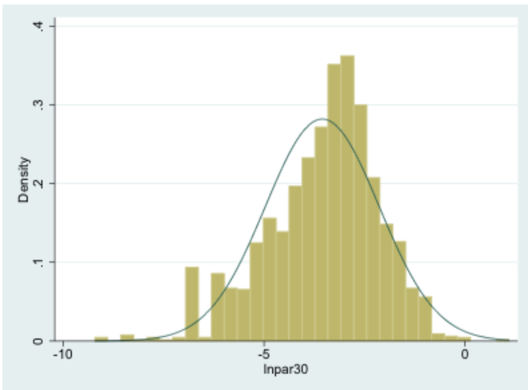
Dependent variable: operexp_assets



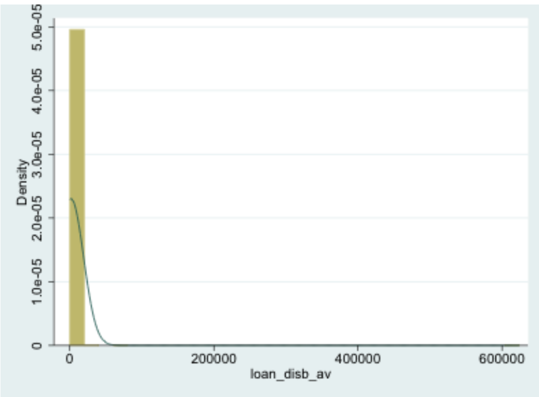


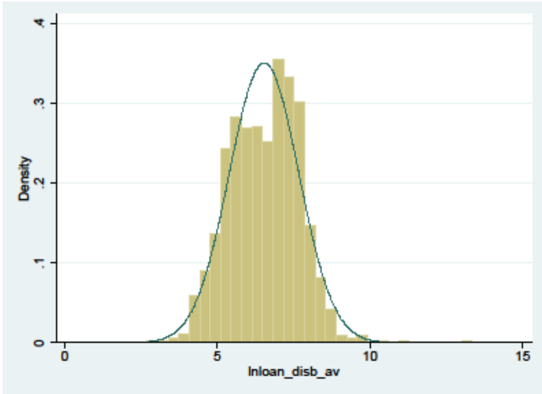
Independent variable: par30



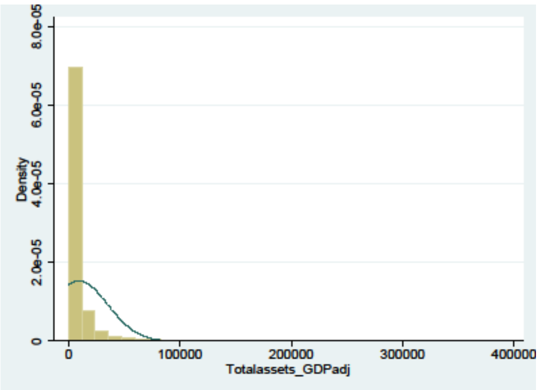


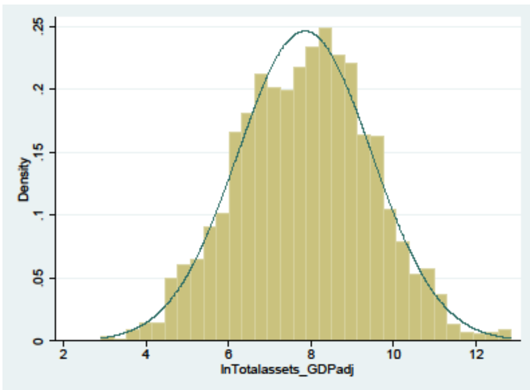
Control variable: loan_disb_av



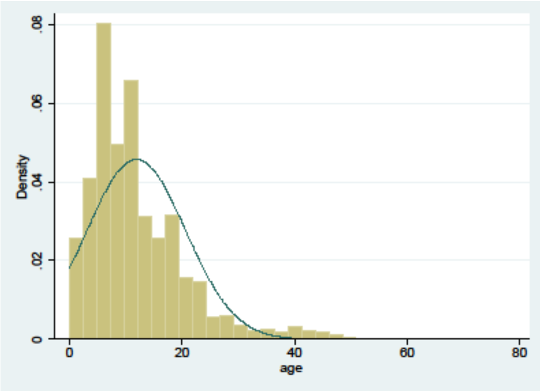


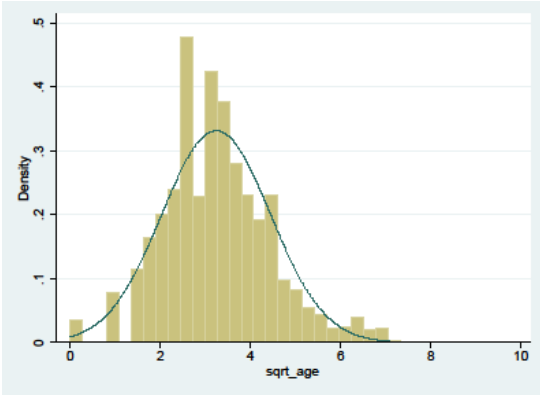
Control variable: lnTotalassets_GDPadj



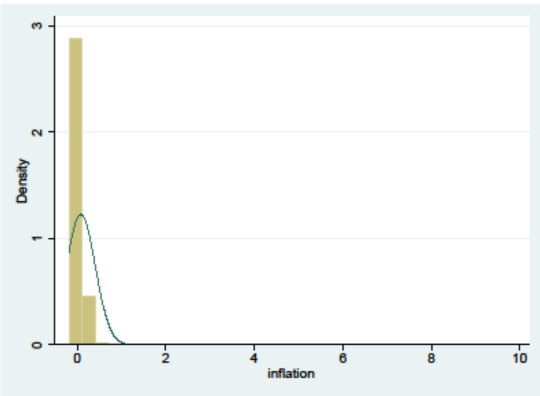


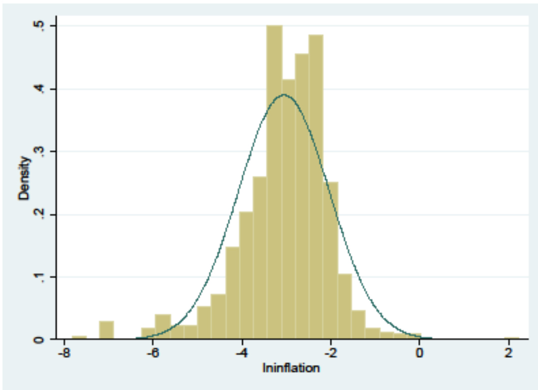
Control variable: sqrt_age



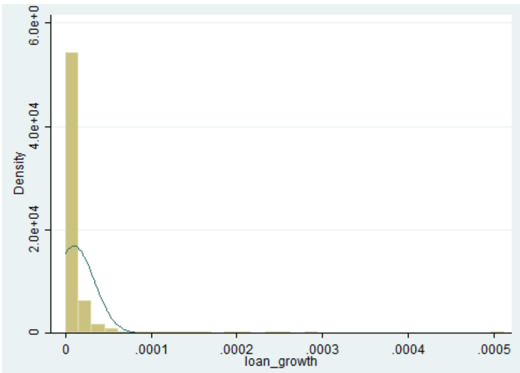


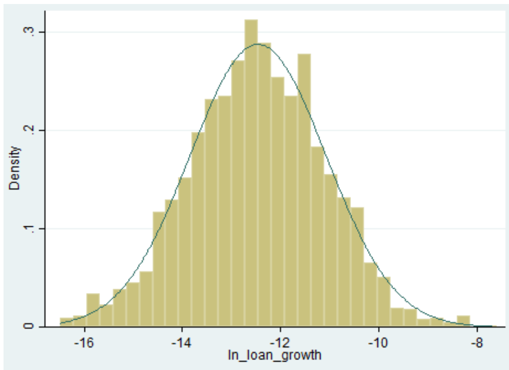
Control variable: lninflation





Control variable: ln_loan_growth





Appendix 2: Regressions using alternative measures

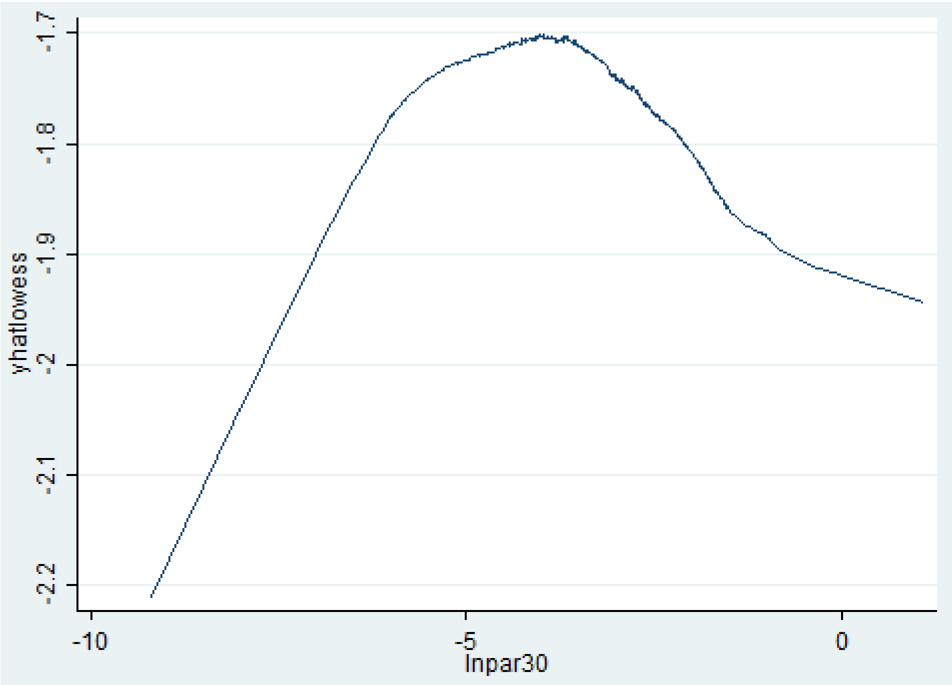
Dependent variable: operating expenses related to portfolio

Independent variable: par30

```
Linear regression                                      Number of obs =       271
                                                       F(7, 263)     =     33.24
                                                       Prob > F      =    0.0000
                                                       R-squared     =    0.4418
                                                       Root MSE      =    .48038
```

lnoperexp_portf	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
lnpar30	.0137606	.0213403	0.64	0.520	-.0282589	.0557801
lnloan_disb_av	-.3230324	.0313551	-10.30	0.000	-.3847714	-.2612934
lnTotalassets_GDPadj	-.0980609	.0201658	-4.86	0.000	-.1377678	-.058354
sqrt_age	-.0417128	.0258753	-1.61	0.108	-.0926619	.0092362
lninflation	.0659037	.0280462	2.35	0.020	.0106799	.1211274
rural	-.5314696	.0880764	-6.03	0.000	-.7048942	-.358045
DM_individ	-.106124	.120591	-0.88	0.380	-.3435707	.1313227
_cons	1.977256	.3023818	6.54	0.000	1.381859	2.572653

Preliminary regression with robust standard errors



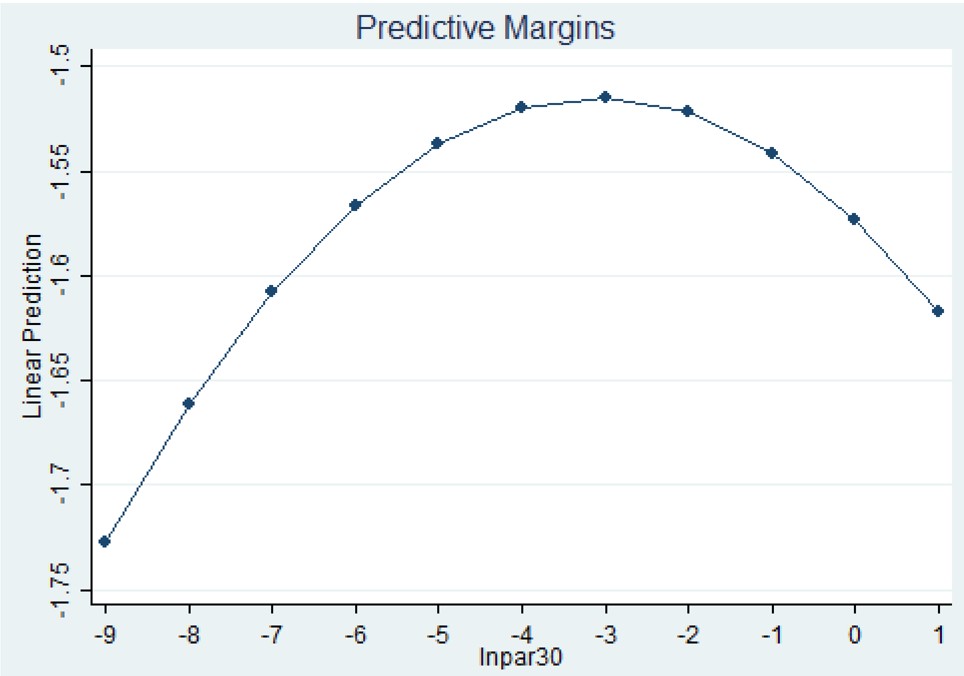
Prediction of bivariate relationship between par30 and operating expenses related to portfolio

Linear regression

Number of obs	=	271
F(8, 262)	=	29.10
Prob > F	=	0.0000
R-squared	=	0.4425
Root MSE	=	.481

lnoperexp_portf	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
lnpar30	-.0377768	.0918835	-0.41	0.681	-.2187009	.1431472
c.lnpar30#c.lnpar30	-.0061037	.0098781	-0.62	0.537	-.0255542	.0133468
lnloan_disb_av	-.3247915	.0314834	-10.32	0.000	-.3867843	-.2627988
lnTotalassets_GDPadj	-.0973802	.0202083	-4.82	0.000	-.1371716	-.0575888
sqrt_age	-.0393406	.0269533	-1.46	0.146	-.0924132	.0137321
lninflation	.0664066	.0281337	2.36	0.019	.0110097	.1218035
rural	-.5321624	.0878226	-6.06	0.000	-.7050904	-.3592344
DM_individ	-.1060904	.1201151	-0.88	0.378	-.3426041	.1304234
_cons	1.882671	.3510223	5.36	0.000	1.191487	2.573854

Regression including quadric term for dependent variable



Predictive margins

```
Fixed-effects (within) regression               Number of obs      =       271
Group variable: case                            Number of groups   =       211

R-sq:                                           Obs per group:
     within  = 0.4186                                         min =          1
     between = 0.0485                                         avg =        1.3
     overall = 0.0562                                         max =          4

                                                F(7,53)            =      5.45
corr(u_i, Xb)  = -0.6283                        Prob > F           =    0.0001

---------------------------------------------------------------------------------------
      lnoperexp_portf |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
----------------------+----------------------------------------------------------------
              lnpar30 |  -.0743089   .0460892    -1.61   0.113    -.1667522    .0181344
         lnloan_disb_av |   .0294079   .1250507     0.24   0.815    -.2214122    .2802279
lnTotalassets_GDPadj |  -.4181571   .1004819    -4.16   0.000    -.6196981    -.216616
             sqrt_age |   .0537929   .1867771     0.29   0.774    -.3208345    .4284202
          lninflation |  -.1267126   .0755537    -1.68   0.099    -.2782542     .024829
                rural |  -.2093738   .2471264    -0.85   0.401    -.7050466    .2862989
           DM_individ |  -.3193475   .4457113    -0.72   0.477    -1.213331    .5746362
                _cons |   1.127591   1.080864     1.04   0.302    -1.040347    3.295529
----------------------+----------------------------------------------------------------
              sigma_u |  .79451038
              sigma_e |  .29398583
                  rho |   .8795726   (fraction of variance due to u_i)
---------------------------------------------------------------------------------------
F test that all u_i=0: F(210, 53) = 3.09                     Prob > F = 0.0000
```

Results from fixed-effects regression

```
Random-effects GLS regression                   Number of obs      =       271
Group variable: case                            Number of groups   =       211

R-sq:                                           Obs per group:
     within  = 0.1845                                         min =          1
     between = 0.4399                                         avg =        1.3
     overall = 0.4362                                         max =          4

                                                Wald chi2(7)       =    170.02
corr(u_i, X)   = 0 (assumed)                    Prob > chi2        =    0.0000

---------------------------------------------------------------------------------------
      lnoperexp_portf |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------------------+----------------------------------------------------------------
              lnpar30 |  -.0087452   .0237898    -0.37   0.713    -.0553723    .0378819
         lnloan_disb_av |  -.3198699   .0333631    -9.59   0.000    -.3852603   -.2544795
lnTotalassets_GDPadj |  -.1143104   .0218854    -5.22   0.000    -.1572051   -.0714158
             sqrt_age |  -.0384954   .0327407    -1.18   0.240     -.102666    .0256753
          lninflation |   .0457927   .0331184     1.38   0.167    -.0191182    .1107035
                rural |  -.4845518   .0851801    -5.69   0.000    -.6515017   -.3176019
           DM_individ |   -.088304   .1159946    -0.76   0.446    -.3156491    .1390412
                _cons |   1.907597    .318233     5.99   0.000     1.283872    2.531323
----------------------+----------------------------------------------------------------
              sigma_u |   .3953784
              sigma_e |  .29398583
                  rho |  .64396686   (fraction of variance due to u_i)
---------------------------------------------------------------------------------------
```

Results from random-effects regression

	Coefficients (b) fixed	(B) random	(b-B) Difference	sqrt(diag(V_b-V_B)) S.E.
lnpar30	-.0743089	-.0087452	-.0655637	.0394748
lnloan_dis~v	.0294079	-.3198699	.3492778	.120518
lnTotalass~j	-.4181571	-.1143104	-.3038466	.0980695
sqrt_age	.0537929	-.0384954	.0922882	.1838851
lninflation	-.1267126	.0457927	-.1725052	.0679083
rural	-.2093738	-.4845518	.275178	.2319823
DM_individ	-.3193475	-.088304	-.2310435	.4303532

```
                  b = consistent under Ho and Ha; obtained from xtreg
   B = inconsistent under Ha, efficient under Ho; obtained from xtreg

Test:  Ho:  difference in coefficients not systematic

              chi2(7) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                      =       24.01
            Prob>chi2 =      0.0011
```

Results from Hausman test

```
Fixed-effects (within) IV regression          Number of obs      =       237
Group variable: case                          Number of groups   =       196

R-sq:                                         Obs per group:
     within  = 0.0221                                        min =         1
     between = 0.1421                                        avg =       1.2
     overall = 0.1432                                        max =         3

                                              Wald chi2(8)       =   1705.47
corr(u_i, Xb)  = -0.5163                      Prob > chi2        =    0.0000
```

operexp_portf	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lnpar30	-.2610812	.2331441	-1.12	0.263	-.7180352	.1958729
lnpar30_sqrt	-.0229378	.0204531	-1.12	0.262	-.0630251	.0171495
lnloan_disb_av	-.0469535	.0523865	-0.90	0.370	-.1496292	.0557222
lnTotalassets_GDPadj	-.0725855	.0411672	-1.76	0.078	-.1532716	.0081007
sqrt_age	-.0434621	.0815425	-0.53	0.594	-.2032825	.1163583
lninflation	.0026565	.0300878	0.09	0.930	-.0563146	.0616275
1.rural	-.0302766	.1054791	-0.29	0.774	-.2370119	.1764587
1.DM_individ	-.0600007	.154601	-0.39	0.698	-.3630131	.2430116
_cons	.7684474	.5879416	1.31	0.191	-.3838971	1.920792
sigma_u	.22043075					
sigma_e	.09899987					
rho	.83214835	(fraction of variance due to u_i)				

```
F  test that all u_i=0:     F(195,33) =      2.90           Prob > F     = 0.0003

Instrumented:   lnpar30
Instruments:    lnpar30_sqrt lnloan_disb_av lnTotalassets_GDPadj sqrt_age lninflation
                1.rural 1.DM_individ lag1par30 lag2par30
```

Results from IV-regression, fixed effects

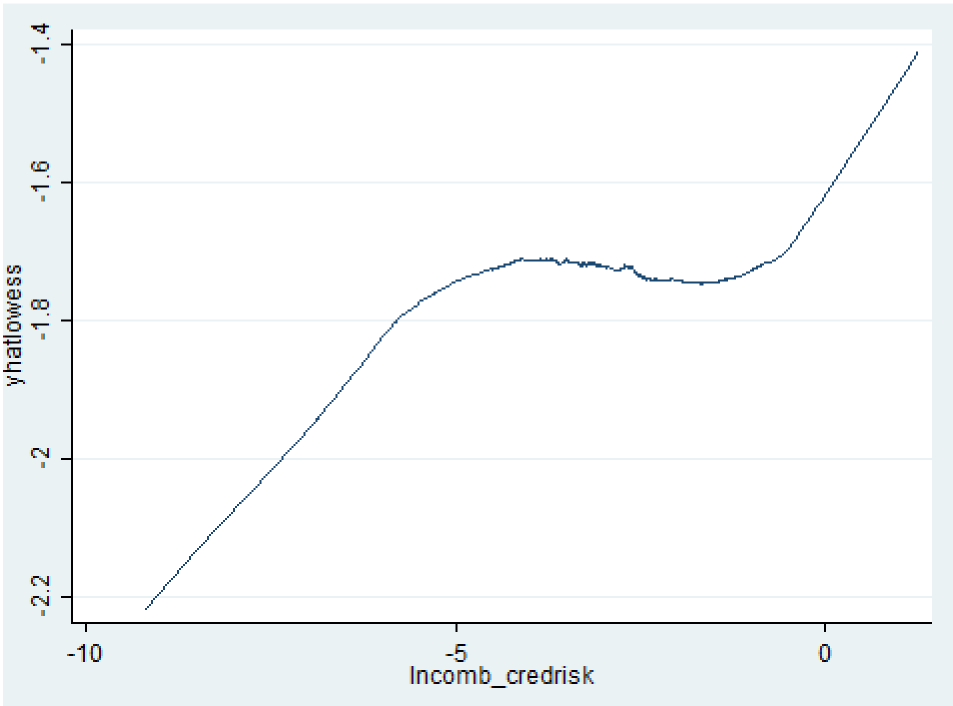
Dependent variable: operating expenses related to assets

Independent variable: combined credit risk

```
Linear regression                               Number of obs     =        693
                                                F(7, 685)         =      65.82
                                                Prob > F          =     0.0000
                                                R-squared         =     0.3455
                                                Root MSE          =     .46915
```

lnoperexp_ass~s	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
lncomb_credrisk	-.0455794	.0159578	-2.86	0.004	-.0769115	-.0142474
lnloan_disb_av	-.2092323	.0206004	-10.16	0.000	-.2496797	-.1687848
lnTotalassets	-.1450356	.0172853	-8.39	0.000	-.1789742	-.111097
sqrt_age	-.0087122	.0177313	-0.49	0.623	-.0435264	.026102
lninflation	.0396025	.016435	2.41	0.016	.0073336	.0718715
rural	-.3673288	.0528248	-6.95	0.000	-.4710468	-.2636107
DM_individ	-.1191037	.0676543	-1.76	0.079	-.2519384	.013731
_cons	2.016086	.248021	8.13	0.000	1.529114	2.503059

Preliminary regression with robust standard errors



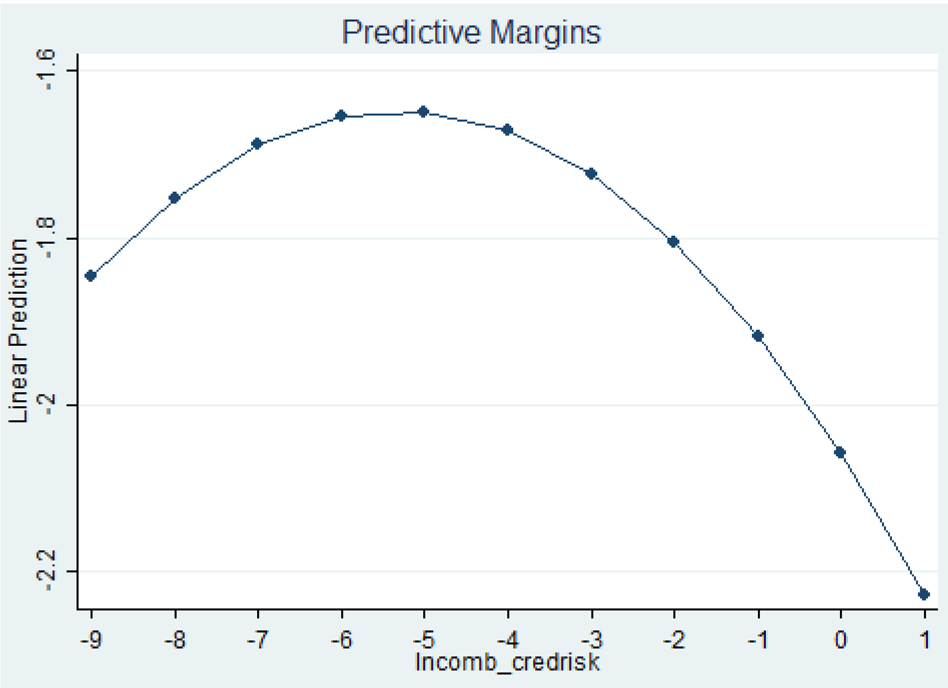
Prediction of bivariate relationship between the combined credit risk and operating expenses related to assets

Note: Regressions and predictive margins below are run using both the quadric and qubic term of the independent variable, as the prediction of the bivariate relationship show a qubic function. Results on predictive margins show a quadric function, which is used in following tests.

```
Linear regression                               Number of obs     =        693
                                                F(8, 684)         =      60.51
                                                Prob > F          =     0.0000
                                                R-squared         =     0.3493
                                                Root MSE          =     .46816
```

lnoperexp_assets	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
lncomb_credrisk	-.1547883	.0560739	-2.76	0.006	-.2648859	-.0446906
c.lncomb_credrisk#c.lncomb_credrisk	-.0145757	.0068554	-2.13	0.034	-.0280359	-.0011156
lnloan_disb_av	-.2112874	.0204082	-10.35	0.000	-.2513576	-.1712171
lnTotalassets	-.1463633	.0171615	-8.53	0.000	-.1800589	-.1126677
sqrt_age	-.0062187	.0179788	-0.35	0.730	-.0415189	.0290815
lninflation	.0434704	.0165416	2.63	0.009	.0109921	.0759488
rural	-.3750218	.0530063	-7.08	0.000	-.4790965	-.2709471
DM_individ	-.1235893	.0678506	-1.82	0.069	-.2568097	.009631
_cons	1.883232	.2640696	7.13	0.000	1.364747	2.401716

Regression including quadric term for dependent variable

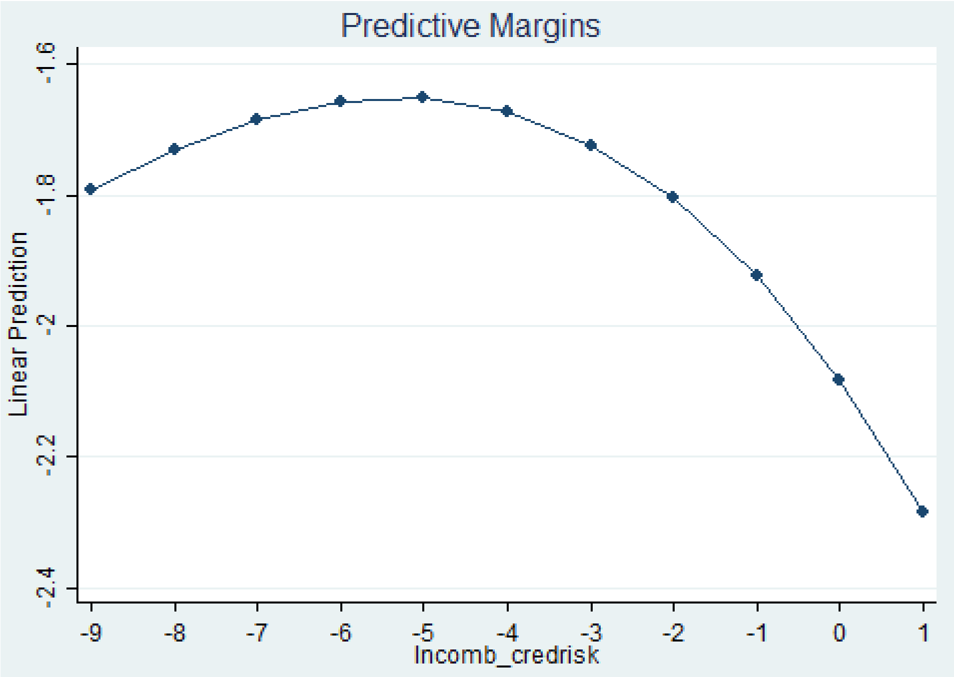


Predictive margins

```
Linear regression                               Number of obs     =        693
                                                F(9, 683)         =      53.99
                                                Prob > F          =     0.0000
                                                R-squared         =     0.3493
                                                Root MSE          =     .46849
```

lnoperexp_assets	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
lncomb_credrisk	-.1791671	.1113628	-1.61	0.108	-.3978217	.0394875
c.lncomb_credrisk#c.lncomb_credrisk	-.0215248	.028019	-0.77	0.443	-.0765386	.033489
c.lncomb_credrisk#c.lncomb_credrisk#c.lncomb_credrisk	-.0005801	.0022072	-0.26	0.793	-.0049139	.0037536
lnloan_disb_av	-.2113373	.0204129	-10.35	0.000	-.2514169	-.1712577
lnTotalassets	-.146286	.0172029	-8.50	0.000	-.1800629	-.112509
sqrt_age	-.0065633	.0180849	-0.36	0.717	-.042072	.0289455
lninflation	.0437458	.01654	2.64	0.008	.0112703	.0762212
rural	-.3752054	.0530585	-7.07	0.000	-.4793828	-.271028
DM_individ	-.1242227	.0680462	-1.83	0.068	-.2578275	.0093821
_cons	1.860503	.2851917	6.52	0.000	1.300545	2.420461

Regression including qubic term for dependent variable



Predictive margins

```
Fixed-effects (within) regression               Number of obs      =       635
Group variable: case                            Number of groups   =       218

R-sq:                                           Obs per group:
     within  = 0.1955                                         min =         1
     between = 0.1484                                         avg =       2.9
     overall = 0.1978                                         max =         8

                                                F(8,409)           =     12.43
corr(u_i, Xb)  = -0.2103                        Prob > F           =    0.0000
```

lnoperexp_assets	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lncomb_credrisk	-.097279	.0437969	-2.22	0.027	-.1833742	-.0111839
lncomb_credrisk_sqrt	-.0081395	.0052	-1.57	0.118	-.0183616	.0020825
lnloan_disb_av	-.0779957	.0395575	-1.97	0.049	-.1557571	-.0002343
lnTotalassets_GDPadj	-.2274414	.0354053	-6.42	0.000	-.2970405	-.1578423
sqrt_age	.0628453	.0550931	1.14	0.255	-.0454558	.1711464
lninflation	.0105007	.014988	0.70	0.484	-.0189624	.0399637
rural	-.0880984	.1323538	-0.67	0.506	-.348277	.1720801
DM_individ	.0413989	.0499164	0.83	0.407	-.0567259	.1395236
_cons	.1527291	.2815409	0.54	0.588	-.4007187	.7061769
sigma_u	.55902764					
sigma_e	.20423388					
rho	.88224527	(fraction of variance due to u_i)				

```
F test that all u_i=0: F(217, 409) = 12.81              Prob > F = 0.0000
```

Results from fixed-effects regression

```
Random-effects GLS regression                   Number of obs      =       635
Group variable: case                            Number of groups   =       218

R-sq:                                           Obs per group:
     within  = 0.1690                                         min =         1
     between = 0.3206                                         avg =       2.9
     overall = 0.3564                                         max =         8

                                                Wald chi2(8)       =    183.26
corr(u_i, X)   = 0 (assumed)                    Prob > chi2        =    0.0000
```

lnoperexp_assets	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lncomb_credrisk	-.092125	.0407971	-2.26	0.024	-.1720859	-.0121641
lncomb_credrisk_sqrt	-.0085112	.0049461	-1.72	0.085	-.0182053	.0011829
lnloan_disb_av	-.1861682	.0249928	-7.45	0.000	-.2351531	-.1371832
lnTotalassets_GDPadj	-.1324395	.0177696	-7.45	0.000	-.1672673	-.0976118
sqrt_age	-.0256961	.0249828	-1.03	0.304	-.0746615	.0232693
lninflation	.0173083	.0140209	1.23	0.217	-.0101722	.0447887
rural	-.2461645	.0757978	-3.25	0.001	-.3947255	-.0976035
DM_individ	.0098348	.0466365	0.21	0.833	-.081571	.1012406
_cons	.4888509	.2061215	2.37	0.018	.0848603	.8928416
sigma_u	.46494878					
sigma_e	.20423388					
rho	.83825792	(fraction of variance due to u_i)				

Results from random-effects regression

	Coefficients (b) fixed	(B) random	(b-B) Difference	sqrt(diag(V_b-V_B)) S.E.
lncomb_cre~k	-.097279	-.092125	-.005154	.0159299
lncomb_cre~t	-.0081395	-.0085112	.0003717	.0016051
lnloan_dis~v	-.0779957	-.1861682	.1081725	.030662
lnTotalass~j	-.2274414	-.1324395	-.0950019	.0306232
sqrt_age	.0628453	-.0256961	.0885414	.0491031
lninflation	.0105007	.0173083	-.0068076	.0052966
rural	-.0880984	-.2461645	.1580661	.1084999
DM_individ	.0413989	.0098348	.031564	.0177958

```
                 b = consistent under Ho and Ha; obtained from xtreg
  B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(8) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =       24.65
                Prob>chi2 =      0.0018
```

Results from Hausman test

```
Fixed-effects (within) IV regression            Number of obs      =       395
Group variable: case                            Number of groups   =       200

R-sq:                                           Obs per group:
     within  = 0.3253                                         min =         1
     between = 0.2053                                         avg =       2.0
     overall = 0.2281                                         max =         6

                                                Wald chi2(8)       =  57275.43
corr(u_i, Xb)  = -0.1553                        Prob > chi2        =    0.0000
```

lnoperexp_assets	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lncomb_credrisk	-.0778432	.0520987	-1.49	0.135	-.1799549	.0242685
lncomb_credrisk_sqrt	-.007366	.0054357	-1.36	0.175	-.0180199	.0032878
lnloan_disb_av	-.0613524	.0496182	-1.24	0.216	-.1586023	.0358975
lnTotalassets_GDPadj	-.1989767	.0363549	-5.47	0.000	-.2702309	-.1277225
sqrt_age	-.0583358	.0676309	-0.86	0.388	-.1908899	.0742183
lninflation	.0082525	.0178396	0.46	0.644	-.0267126	.0432175
1.rural	-.0488076	.1231098	-0.40	0.692	-.2900983	.1924831
1.DM_individ	.0094488	.0523294	0.18	0.857	-.0931149	.1120125
_cons	.2995296	.3418913	0.88	0.381	-.3705651	.9696243
sigma_u	.50422307					
sigma_e	.14611218					
rho	.92253422	(fraction of variance due to u_i)				

```
F  test that all u_i=0:     F(199,187) =     15.41          Prob > F     = 0.0000

Instrumented:   lncomb_credrisk
Instruments:    lncomb_credrisk_sqrt lnloan_disb_av lnTotalassets_GDPadj sqrt_age
                lninflation 1.rural 1.DM_individ lncomb_credrisk
```

Results from IV-regression, fixed effects

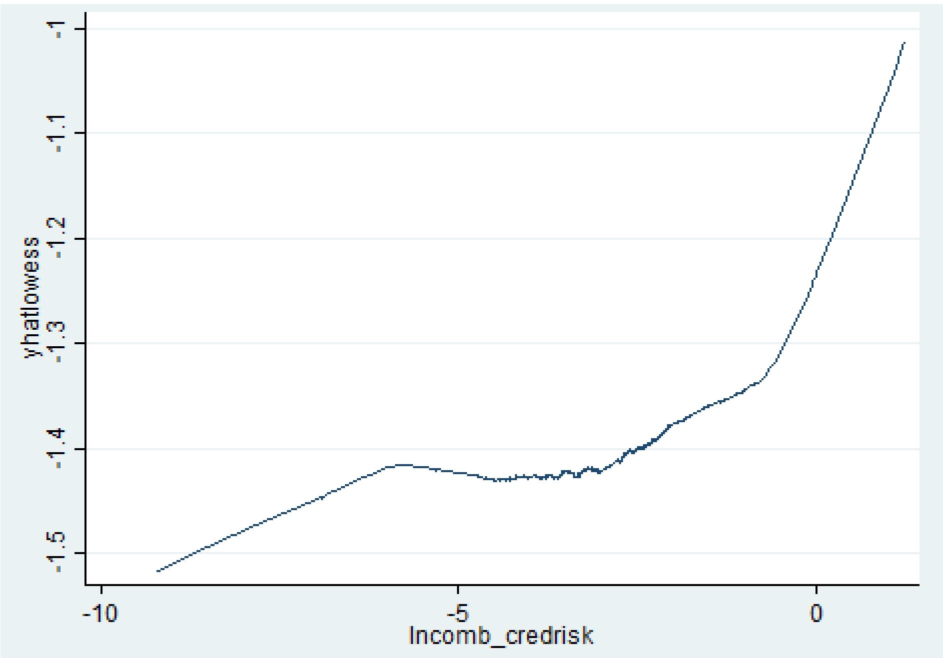
Dependent variable: operating expenses related to portfolio

Independent variable: combined credit risk

Linear regression			
	Number of obs	=	706
	F(7, 698)	=	108.80
	Prob > F	=	0.0000
	R-squared	=	0.4678
	Root MSE	=	.47428

lnoperexp_portf	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
lncomb_credrisk	.0334251	.0140943	2.37	0.018	.0057528	.0610973
lnloan_disb_av	-.3370255	.0172558	-19.53	0.000	-.3709049	-.303146
lnTotalassets_GDPadj	-.1177161	.0128815	-9.14	0.000	-.1430072	-.0924251
sqrt_age	-.0576284	.0178403	-3.23	0.001	-.0926556	-.0226012
lninflation	.0425953	.0157286	2.71	0.007	.0117143	.0734764
rural	-.3395798	.0498876	-6.81	0.000	-.4375276	-.241632
DM_individ	-.1952349	.0680246	-2.87	0.004	-.3287923	-.0616774
_cons	2.351242	.1762535	13.34	0.000	2.005192	2.697293

Preliminary regression with robust standard errors

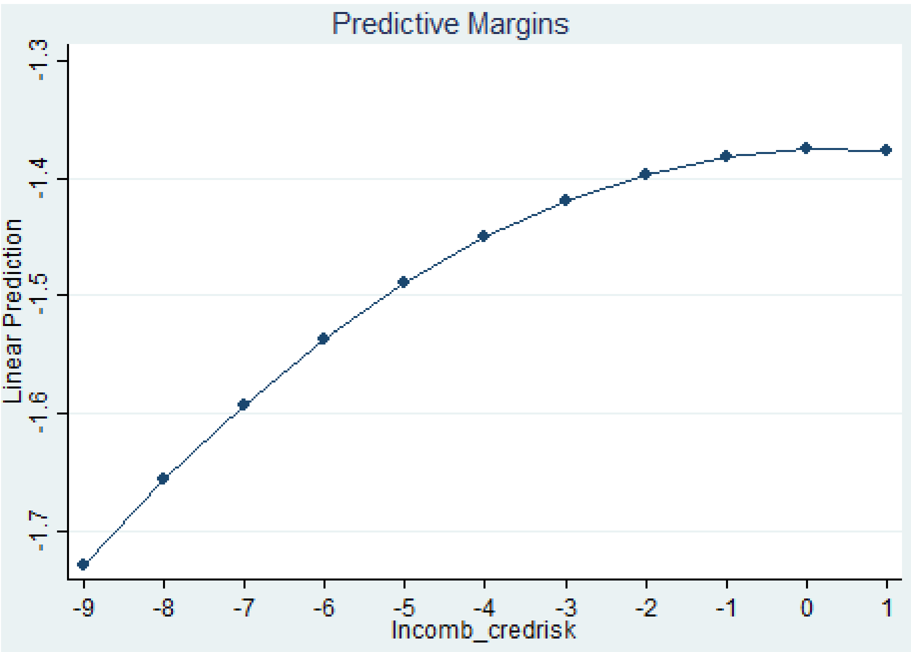


Prediction of bivariate relationship between the combined credit risk and operating expenses related to portfolio

Linear regression			
	Number of obs	=	706
	F(8, 697)	=	98.51
	Prob > F	=	0.0000
	R-squared	=	0.4680
	Root MSE	=	.47452

lnoperexp_portf	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
lncomb_credrisk	.0023598	.0598554	0.04	0.969	-.1151587	.1198783
c.lncomb_credrisk#c.lncomb_credrisk	-.004097	.0073194	-0.56	0.576	-.0184678	.0102738
lnloan_disb_av	-.3376051	.0172327	-19.59	0.000	-.3714394	-.3037709
lnTotalassets_GDPadj	-.1181521	.0127207	-9.29	0.000	-.1431276	-.0931766
sqrt_age	-.0569163	.0179082	-3.18	0.002	-.0920768	-.0217558
lninflation	.043106	.0158005	2.73	0.007	.0120838	.0741283
rural	-.3416938	.0502221	-6.80	0.000	-.4402986	-.2430891
DM_individ	-.1967709	.0679205	-2.90	0.004	-.3301242	-.0634177
_cons	2.309222	.2069344	11.16	0.000	1.902932	2.715511

Regression including quadric term for dependent variable



Predictive margins

```
Fixed-effects (within) regression               Number of obs      =       706
Group variable: case                            Number of groups   =       219

R-sq:                                           Obs per group:
     within  = 0.2938                                         min =          1
     between = 0.1527                                         avg =        3.2
     overall = 0.2314                                         max =          8

                                                F(8,479)           =     24.92
corr(u_i, Xb)  = -0.2491                        Prob > F           =    0.0000
```

lnoperexp_portf	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lncomb_credrisk	-.0496887	.0421729	-1.18	0.239	-.1325554	.0331779
lncomb_credrisk_sqrt	-.0039162	.0051062	-0.77	0.443	-.0139494	.006117
lnloan_disb_av	-.0929835	.036482	-2.55	0.011	-.164668	-.021299
lnTotalassets_GDPadj	-.2706473	.0346504	-7.81	0.000	-.3387328	-.2025618
sqrt_age	-.0473428	.0494876	-0.96	0.339	-.1445823	.0498968
lninflation	-.0135919	.0147459	-0.92	0.357	-.0425666	.0153829
rural	-.140221	.1436716	-0.98	0.330	-.4225256	.1420835
DM_individ	.0545544	.0515042	1.06	0.290	-.0466478	.1557565
_cons	1.258736	.2584405	4.87	0.000	.7509191	1.766553
sigma_u	.62523733					
sigma_e	.22259946					
rho	.88750592	(fraction of variance due to u_i)				

```
F test that all u_i=0: F(218, 479) = 12.33                 Prob > F = 0.0000
```

Results from fixed-effects regression

```
Random-effects GLS regression                   Number of obs      =       706
Group variable: case                            Number of groups   =       219

R-sq:                                           Obs per group:
     within  = 0.2465                                         min =          1
     between = 0.4042                                         avg =        3.2
     overall = 0.4326                                         max =          8

                                                Wald chi2(8)       =    307.25
corr(u_i, X)   = 0 (assumed)                    Prob > chi2        =    0.0000
```

lnoperexp_portf	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lncomb_credrisk	-.0017303	.0406367	-0.04	0.966	-.0813766	.0779161
lncomb_credrisk_sqrt	-.0005005	.0049744	-0.10	0.920	-.0102502	.0092491
lnloan_disb_av	-.2628243	.0246943	-10.64	0.000	-.3112243	-.2144243
lnTotalassets_GDPadj	-.1483403	.0179817	-8.25	0.000	-.1835837	-.1130969
sqrt_age	-.0726103	.0246235	-2.95	0.003	-.1208715	-.0243491
lninflation	.0008329	.0141969	0.06	0.953	-.0269925	.0286582
rural	-.2955625	.0788411	-3.75	0.000	-.4500882	-.1410367
DM_individ	-.0038997	.0489355	-0.08	0.936	-.0998115	.0920121
_cons	1.733998	.2019961	8.58	0.000	1.338093	2.129903
sigma_u	.45663981					
sigma_e	.22259946					
rho	.80799614	(fraction of variance due to u_i)				

Results from random-effects regression

	Coefficients (b) fixed	(B) random	(b-B) Difference	sqrt(diag(V_b-V_B)) S.E.
lncomb_cre~k	-.0496887	-.0017303	-.0479585	.0112789
lncomb_cre~t	-.0039162	-.0005005	-.0034157	.0011525
lnloan_dis~v	-.0929835	-.2628243	.1698408	.0268538
lnTotalass~j	-.2706473	-.1483403	-.122307	.0296194
sqrt_age	-.0473428	-.0726103	.0252676	.0429267
lninflation	-.0135919	.0008329	-.0144248	.0039865
rural	-.140221	-.2955625	.1553414	.1201067
DM_individ	.0545544	-.0038997	.0584541	.0160626

b = consistent under Ho and Ha; obtained from xtreg
B = inconsistent under Ha, efficient under Ho; obtained from xtreg

Test: Ho: difference in coefficients not systematic

chi2(8) = (b-B)'[(V_b-V_B)^(-1)](b-B)
= 67.53
Prob>chi2 = 0.0000

Results from Hausman test

```
Fixed-effects (within) IV regression           Number of obs      =       399
Group variable: case                           Number of groups   =       202

R-sq:                                          Obs per group:
     within  = 0.3197                                        min =          1
     between = 0.0883                                        avg =        2.0
     overall = 0.1115                                        max =          6

                                               Wald chi2(8)       =  23434.33
corr(u_i, Xb)  = -0.3725                       Prob > chi2        =    0.0000
```

lnoperexp_portf	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lncomb_credrisk	-.0611695	.0679067	-0.90	0.368	-.1942642	.0719252
lncomb_credrisk_sqrt	-.0027973	.0071001	-0.39	0.694	-.0167132	.0111186
lnloan_disb_av	-.0116225	.0659439	-0.18	0.860	-.1408703	.1176252
lnTotalassets_GDPadj	-.2271791	.0480164	-4.73	0.000	-.3212896	-.1330686
sqrt_age	-.1685946	.0888771	-1.90	0.058	-.3427904	.0056013
lninflation	-.0310698	.0233431	-1.33	0.183	-.0768215	.0146819
1.rural	-.0155782	.164148	-0.09	0.924	-.3373024	.306146
1.DM_individ	.0643357	.069654	0.92	0.356	-.0721836	.200855
_cons	.6610466	.4558699	1.45	0.147	-.2324419	1.554535
sigma_u	.62787097					
sigma_e	.19483568					
rho	.91216464	(fraction of variance due to u_i)				

F test that all u_i=0: F(201,189) = 8.73 Prob > F = 0.0000

Instrumented: lncomb_credrisk
Instruments: lncomb_credrisk_sqrt lnloan_disb_av lnTotalassets_GDPadj sqrt_age lninflation 1.rural 1.DM_individ lncomb_credrisk

Results from IV-regression, fixed effects

Appendix 3: Independence of error terms (autocorrelation)

	lnoper~s	L. lnoper~s	lnpar30	L. lnpar30	lnloan~v	L. lnloan~v	lnTota~j	L. lnTota~j	sqrt_age	L. sqrt_age	lninfl~n	L. lninfl~n	rural	L. rural
lnoperexp_~s														
--.	1.0000													
L1.	0.9097	1.0000												
lnpar30														
--.	-0.1400	-0.1799	1.0000											
L1.	-0.1155	-0.1598	0.7969	1.0000										
lnloan_dis~v														
--.	-0.4963	-0.4735	0.1522	0.1258	1.0000									
L1.	-0.5193	-0.5036	0.1833	0.1693	0.9640	1.0000								
lnTotalass~j														
--.	-0.4251	-0.3591	-0.1655	-0.2141	0.1451	0.1230	1.0000							
L1.	-0.4365	-0.3791	-0.1220	-0.1857	0.1548	0.1414	0.9844	1.0000						
sqrt_age														
--.	-0.1375	-0.1093	0.4386	0.3928	0.0637	0.0808	0.0190	0.0373	1.0000					
L1.	-0.1401	-0.1076	0.4373	0.3923	0.0591	0.0776	0.0205	0.0406	0.9993	1.0000				
lninflation														
--.	0.1033	0.0838	0.0665	0.0398	-0.0162	-0.0285	0.0461	0.0551	0.0365	0.0352	1.0000			
L1.	0.0575	0.0550	0.1405	0.0980	-0.0440	-0.0519	-0.0236	-0.0088	0.0323	0.0307	0.5905	1.0000		
rural														
--.	-0.1459	-0.1232	-0.1034	-0.0916	-0.0482	-0.0543	0.0339	0.0150	-0.0269	-0.0305	-0.0857	-0.0847	1.0000	
L1.	-0.1507	-0.1340	-0.1133	-0.1077	-0.0489	-0.0582	0.0401	0.0258	-0.0315	-0.0347	-0.0714	-0.0755	0.9752	1.0000

Appendix 4: Fixed and Random effects

Fixed-effects (within) regression — Number of obs = 266
Group variable: case — Number of groups = 210

R-sq: — Obs per group:
within = 0.4191 — min = 1
between = 0.1184 — avg = 1.3
overall = 0.1086 — max = 4

F(8,48) = 4.33
corr(u_i, Xb) = -0.4684 — Prob > F = 0.0006

lnoperexp_assets	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lnpar30	.0026245	.1159667	0.02	0.982	-.230542	.2357911
lnpar30_sqrt	-.0027884	.0105464	-0.26	0.793	-.0239934	.0184166
lnloan_disb_av	-.0757595	.087846	-0.86	0.393	-.2523857	.1008668
lnTotalassets_GDPadj	-.2847474	.0691391	-4.12	0.000	-.4237609	-.145734
sqrt_age	.11472	.135431	0.85	0.401	-.1575822	.3870222
lninflation	.0507573	.0564977	0.90	0.373	-.0628388	.1643535
rural	-.1315013	.1679229	-0.78	0.437	-.469133	.2061303
DM_individ	-.0445337	.3012973	-0.15	0.883	-.6503326	.5612652
_cons	.8754812	.7740489	1.13	0.264	-.6808484	2.431811
sigma_u	.6081298					
sigma_e	.1985164					
rho	.90370036	(fraction of variance due to u_i)				

F test that all u_i=0: F(209, 48) = 6.23 — Prob > F = 0.0000

Random-effects GLS regression — Number of obs = 266
Group variable: case — Number of groups = 210

R-sq: — Obs per group:
within = 0.2998 — min = 1
between = 0.3688 — avg = 1.3
overall = 0.3610 — max = 4

Wald chi2(8) = 140.22
corr(u_i, X) = 0 (assumed) — Prob > chi2 = 0.0000

lnoperexp_assets	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lnpar30	-.1403935	.0717609	-1.96	0.050	-.2810424	.0002553
lnpar30_sqrt	-.0142898	.0076286	-1.87	0.061	-.0292415	.000662
lnloan_disb_av	-.2340021	.0304219	-7.69	0.000	-.2936279	-.1743762
lnTotalassets_GDPadj	-.118702	.0203036	-5.85	0.000	-.1584964	-.0789076
sqrt_age	-.0406042	.0310122	-1.31	0.190	-.1013869	.0201786
lninflation	.0608863	.0294573	2.07	0.039	.0031511	.1186215
rural	-.3733561	.0769881	-4.85	0.000	-.52425	-.2224621
DM_individ	-.0514372	.1058604	-0.49	0.627	-.2589198	.1560455
_cons	.835727	.3198523	2.61	0.009	.2088279	1.462626
sigma_u	.4334776					
sigma_e	.1985164					
rho	.8266311	(fraction of variance due to u_i)				

Appendix 5: Reflection note

1.0 Introduction

This reflective note will shortly present the main theme and findings of the thesis, and then continue by identifying how the thesis topic relates to broader international trends, innovation and responsibility.

2.0 Summary of thesis findings

This thesis is based on contract and agency theory, theory on moral hazard, credit risk and Berger & De Youngs (1997) study on cost efficiency. It argues that there is reason to believe that the level of default on loans in microfinance institutions can both positively and negatively affect the operating costs of the institutions. Existing theory and research makes good arguments for both a positive and negative influence of the default level on operating costs. These arguments set the basis for investigating the actual relationship between the two variables. The results of the study show that there is a bivariate quadric relationship between default on loans and operating costs, but this is not supported in the multivariate regression analysis. The consequence of this is that we cannot conclude that the relationship between default on loans and operating costs is non-linear, and we do not succeed in illustrating a general functional form for the relationship. That means that we cannot find any patterns that indicate that a certain level of default will leave the institutions with the ability to minimize operating costs. Further studies are suggested on this topic as the bivariate results suggests that there might in fact be a non-linear relationship, and theory suggests that operating costs, and ultimately lending rates, can be reduced if default rates are optimised.

3.0 Internationalization

Microfinance institutions and other charitable organisations are in many cases reliant on donations in order to carry out their work. These can be donations from governments, large corporations, donor organisations or individuals. The donations are often given across borders, and not only to charitable organisations, but also to developing countries. A global economy has been established as far as donations go. Although microfinance institutions to a larger degree have become self-sufficient, the industry still depends upon these donations. This thesis argues that the industry face a need for lower operating costs in order to lower the interest rate they offer to their clients. But another interesting way of thought is that if microfinance institutions are able to lower their operating costs, they could potentially increase their earnings. The core goal of microfinance institutions is to offer financing to the poor, but another important part of the industry is to do so in a financially sustainable way. Arguably, microfinance institutions will want to lower their interest rates when operating costs decrease, but perhaps they can split the savings from lower operating cost into covering reduced earnings as a result of lower interest charged as well as taking some (more) profit? The idea is that if the institutions are financially sustainable they will be less dependent on donations and become more self-sufficient. The institutions that are able to achieve good profits would perhaps even be able to attract investors, rather than donors. The upside from bringing in investors is that the institutions to a larger degree will be forced to put extra thought into how to achieve efficient and effective operations. Surely, many microfinance institutions have a strong focus on this already, but it is reasonable to assume that some rest comfortably on the fact that they are receiving funding through donations even if operations are not great. Another positive effect of attracting investors is that the chances are that they will be able to offer more funding than the microfinance institutions can achieve through donations. These extra resources can be used for research and development or other measures that would further improve efficiency. Microfinance institutions could possibly end up finding themselves in a blooming circle of opportunities by taking measures to reduce operating costs. Furthermore, the industry is today facing an increasing demand for international funding as many of the countries the microfinance institutions operate in are developing countries with poor or no possibility to offer donations or support to the institutions. In such a global economy, the competition between the microfinance institutions, charitable organisations and developing countries for donation is great and increasing. It is rational to assume that the competition will further increase in the coming years, as the microfinance industry still experience great growth. The more institutions, the harder the competition. In order to “win” funding by donors the institutions will have to distinguish themselves from other organisations. A great way to do so is by lowering their operating costs. International donors and investors are likely to be more concerned with financial sustainability than the loacals, as they often come from well-developed countries and economies and are used to being able to set certain expectations for their investments/donations. They will want to put their money into well-driven and functioning organisations that are achieving their social goal of outreach as well as being financial sustainability. The way I see it, microfinance institutions can only win on lowering their operating costs. That be if they do it to increase outreach, profitability or sustainability. This is especially true when they operate and compete in an international economy. As concluded in this thesis, further research is needed to determine the relationship between the default on loans and operating costs in microfinance institutions, but I do believe that there is potential for lowering operating costs by investigating this relationship further. Additional drivers of the operating costs should be explored, in order for the microfinance industry to continue to thrive in the global economy that includes countries all over the world.

4.0 Innovation

As explained in this thesis, microfinance institutions serve a social mission of outreach to the poor. This counts for short of half of the world's population. Evidently, microfinance institutions carry great responsibilities in the work they do, and the outreach is very important in order to help more people. To my knowledge, no previous studies have been conducted on the effect of default on loans on operating costs in microfinance institutions. There have been studies on this in commercial banks, but for some reason no one has looked into the microfinance industry. Because the industry over time has become so concerned with lowering risk, the default rates are very low. This is of course not for free. Microfinance institutions spend great resources on monitoring and collection practice, which in turn drives up the operating costs. This thesis investigates whether it is possible to lower operating costs by allowing for more default, but the results are inconclusive. However, the goal is to reduce operating costs, not to increase default. What if it is possible to let the low default levels remain, but still reduce the amount of resources spent on monitoring, control and collection of repayment? One possible way to deal with this challenge is to embrace technology. We have seen great innovation in commercial banking through the use of technology, and would expect to see the same positive effects in the microfinance industry. There is great potential when it comes to technology, and by the use of innovation the microfinance industry could possibly close the gap that has emerged in this respect. The microfinance industry is far behind the commercial banking industry when it comes to technology and need to address this deviation. The challenge is of course that in many of the poorest areas the customers do not have access to power, let alone Internet. The microfinance institutions would need to find innovative ways to reach out to customers with their technology. One way to do so is to set up Internet stations in public areas (much like ATMs) where clients can go to register information, send loan applications, check their debt, etc. This should be connected to a regular ATM and deposit machine in areas where cash is commonly used, so that the clients can also make repayments or receive their loans via the station. Another way to innovate the industry is to implement technological solutions where possible, e.g. in urban and semi-rural areas. Wherever the microfinance institutions can reduce the use of resources is a step in the right direction, as they can allocate the resources elsewhere. A segmentation of the market into geographical areas is thus one way to innovate the lending and collection process.

5.0 Responsibility

Over the past decade the microfinance industry has received criticism for having too harsh collection practise and for neglecting their social mission to serve the poor. Critics argue that the microfinance institutions are exploiting the poor by demanding such high interest rates. As stated in the introduction of this thesis, it is a paradox that the poorest people pay the highest interest rates. Microfinance institutions face an ethical responsibility when it comes to the quality of the services they provide. Though microfinance institutions operate as banks, they are in fact based on a charity concept. The issue of responsibility recur throughout this thesis by pointing out the need for lower operating costs in order to provide lower interest rates to the microfinance clients. Contrary to assertions by critics, research show that microfinance institutions are not being greedy, but that the interest rates offered to the customers are mainly comprised of operating costs, cost of loans and loan loss. High costs and low margins have been pointed out as the main problem of the industry. This thesis encounters the question of responsibility in the microfinance industry by looking into a possible way for the institutions to offer lower interest to their customers. In the future I hope to see that microfinance institutions do not only offer banking to the poor, but that they do so at the same terms as are offered by commercial banks to better-off customers.